

## Durham Research Online

---

### Deposited in DRO:

29 July 2015

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Pontikes, E. G. and Hannan, M. T. (2014) 'An ecology of social categories.', *Sociological science.*, 1 . pp. 311-343.

### Further information on publisher's website:

<http://dx.doi.org/10.15195/v1.a20>

### Publisher's copyright statement:

© 2014 Pontikes and Hannan. This open-access article has been published and distributed under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited.

### Additional information:

---

### Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# An Ecology of Social Categories

Elizabeth G. Pontikes,<sup>a</sup> Michael T. Hannan<sup>b</sup>


a) University of Chicago; b) Stanford University

**Abstract:** This article proposes that meaningful social classification emerges from an ecological dynamic that operates in two planes: feature space and label space. It takes a dynamic view of classification, allowing objects' movements in both spaces to change the meaning of social categories. The first part of the theory argues that agents assign labels to objects based on perceptions of their similarities to existing members of a category. The second part of the theory shows that an object's perceived similarity to members of other categories reduces its typicality in a focal category. This means that for categories with a high degree of overlap with other categories in label space (lenient categories), the link between feature-based similarities and labeling weakens. The findings suggest that social classification will likely evolve to contain both constraining and lenient categories. The theory implies that this process is self-reinforcing, so that constraining categories become more constraining, whereas lenient categories become more lenient.

**Keywords:** categories; features; labels; leniency; patents; software

**Editor(s):** Olav Sorenson; **Received:** April 15, 2014; **Accepted:** May 28, 2014; **Published:** August 18, 2014

**Citation:** Pontikes, Elizabeth G. and Michael T. Hannan. 2014. "An Ecology of Social Categories." *Sociological Science* 1: 311-343. DOI: 10.15195/v1.a20

**Copyright:** © 2014 Pontikes and Hannan. This open-access article has been published and distributed under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited. 

CATEGORIES structure social domains. They locate actors in comparative sets and delineate standards for evaluation. Social agreement about what constitutes membership in categories imposes constraint that induces actors to conform to categorical types. This is an instance of a general dynamic that has long interested sociologists: social structures constrain individual actors, then widespread conformity leads to the replication of the structures.

Research documents the homogenizing effects of social categories. Organizations adopt similar procedures, architectures, and even identities, not only because the practices are effective but also to conform to taken-for-granted ideas about how to organize. What drives this dynamic is that objects have less appeal and receive fewer rewards when they do not fit into an accepted category (Zuckerman 1999; Hsu 2006; Hsu et al. 2009; Leung and Sharkey 2014). In this view, complying with categorical imperatives provides benefits. As a result, actors conform to categorical standards and existing category codes get replicated. This rendering depicts agents as navigating a static category structure when deciding how to classify or make sense of objects and actors.

There is reason to question whether categorical structures tend toward stasis. The vibrant lines of research on categories and markets build on the view that audiences and market intermediaries create categories. Audience members come to agree on labels and codes as they seek to interpret action in a domain. Previous research has focused on how agreed-upon categorical meanings affect perceptions and evaluations. But the converse also holds: the characteristics of objects that get categorized shape meanings. For example, when telephone technology migrated from rotary-dial to touch-tone, the meaning of <telephone> evolved to include the new technology.<sup>1</sup> More generally, when an object changes its characteristics or gets reclassified, this potentially changes the meaning of one or more categories. This means that social categories are less durable than we generally assume.

The question of whether social structures affect the behavior of individual actors, or whether individual behaviors shape social structure, is pertinent to many areas in sociological research. Scholars studying institutions and organizational forms maintain that social structures are homog-

<sup>1</sup>Throughout we mark terms from the language of the domain, the so-called object language, with angle brackets.

enizing and self-replicating while acknowledging that these structures—and the meanings they invoke—frequently change. This tension has been extensively discussed as an important unresolved issue in organizational sociology (Romanelli 1991; Clemens and Cook 1999). That social structures are simultaneously rigid enough to induce widespread conformity, and malleable enough to change as a result of nonconformist behavior, presents an underlying tension in the literature.

Recent work on classification sidesteps this issue by assuming that category boundaries are policed by market intermediaries such as stock market analysts (Zuckerman 1999), credit agencies (Ruef and Patterson 2009), critics (Hsu 2006; Rao et al. 2003; Negro et al. 2010; Negro and Leung 2013), and government agencies (Ruef 2000). But what has historically interested sociologists in categorization is that macro-level structures emerge through uncoordinated individual beliefs about social meaning embedded in objects or concepts (Berger and Luckmann 1967; Durkheim 1931; Meyer and Rowan 1977). For instance, sociological studies of markets show that stable market structure can arise when organizations independently take market positions in response to positions taken by competitive rivals (White 1981; Porac et al. 1995). These ideas also underpin recent work that models the emergence of categories on the basis of audience members coming to agreement about schemas that capture the bases of similarity among objects (Hannan et al. 2007).

We suggest that meaningful classification can emerge even in an evolving environment comprising loosely coordinated actors. This occurs when people observe current categorization, make similarity judgments based on objects' features, classify them accordingly, and share their classifications. If the process is not coordinated, some objects might get reclassified frequently. This can change the meanings of categories, which then alters how objects with stable categorizations get interpreted. It might seem that this type of process would lead to chaos, with meanings and memberships changing too rapidly to coalesce into a coherent system of categories. In contrast, we show that classification can emerge even when objects change features and categorical meanings evolve.

We are struck that this will produce an ecological dynamic for social categorization that operates across two planes: *feature space* and *label space*. Feature space locates objects by their feature values. Label space contains labels for categories. The two planes are connected when agents apply a label to a profile of feature values. When an audience reaches consensus about the link between a label and a set of features, the label denotes a category.

Associations between category labels and features are straightforward in a static world. But categorization becomes more complex when we consider that the social meaning of categories can change. For example, when many category members alter their characteristics, this shifts how people conceive of the category. This means a focal category member who was once typical of that category might become atypical. Such an outcome characterizes obsolescence processes. A scientist who still teaches and researches what she learned decades ago might have been a highly typical member of her discipline as a new graduate but very atypical today. More generally, changes in feature values by some members can affect categorization for others.

A similar dynamic plays out for labeling. If atypical entities are assigned membership in a category—as when producers of processed foods start being labeled as <natural>—then the category will become less crisp. As a result, existing members may no longer be considered typical of the category. For instance, executives of organizations with nanotechnology capabilities concluded that their organizations should not be included in the <nanotechnology> category when the label was used so broadly that its underlying meaning no longer reflected their companies' capabilities (Granqvist et al. 2013).

We study these dynamics by building an explicit theoretical model of the coupling of feature space and label space. Our model relies on two processes, both based on similarity judgments. The first involves the similarity of a newly observed object to the current membership of a category. We argue (and show empirically) that *feature-based similarities* strongly influence label assignments even as actors change positions in both spaces. As objects move, conceptions of a category can change, and so can its membership.

But a link between features and labels persists. In the <telephone> example, the sets of features associated with this category have evolved from rotary dial to touch-tone to touch-screens and small hand-held electronic devices. The specific characteristics associated with the concept have changed, but it still retains a well-defined meaning. The second process underlying our model concerns the *similarity of a focal category to other categories*. Our theory and empirical findings imply that extensive overlap in category memberships weakens the link between feature values and label assignments. This can be seen in the example of <nanotechnology>, which evolved to overlap with a number of scientific disciplines including organic chemistry, molecular biology, semiconductor physics, and so on. Not only does the meaning of this category evolve, so too does the strength of the link between the label and set of features expected of category members.

In addressing these issues we employ a formal language proposed by Hannan et al. (2007) and further elaborated by Hsu et al. (2009) and Kovács and Hannan (2011). We think a formal theory is needed because these notions can be slippery and erroneous reasoning can easily result. Formal languages work against tendencies in natural language to skip steps and misidentify required assumptions, which can lead to imprecise conclusions.

## Category Labels in a Domain

Labels play a central role in our analysis because they denote concepts, mental representations of similarity and difference (Murphy 2002). Although such representations need not be paired with labels, they often are. In the case of social classification, we generally expect concepts to be labeled, otherwise people could not easily communicate about what they represent. Sociological analysis focuses on categories, where a label not only represents a concept for a particular individual but also conveys broad agreement about the pairing of a concept with a label, which gives the label a social meaning. Because people use labels to communicate about concepts and categories, much useful analysis can be done with primary reference to their labels. For this reason we can simplify the language of our argument

by restricting the term “label” to refer to tags that are paired with meanings. We often refer to labels as shorthand for the concepts/categories that they designate.

We construct a theory for the general case in which agents, as audience, classify producers. Audience and producer are paired roles, and in many contexts the same actors play both roles.<sup>2</sup> That is, producers can be audience to each other. We do not specify a particular kind of audience because our assumptions plausibly describe processes that apply for many kinds of agents playing an audience role, including critics, consumers, rivals, and the producers themselves.

We build our argument at the level of an individual person as an audience member. This focus cuts a layer of complexity that arises because different audiences have different perspectives on classification (Pontikes 2012; Hsu and Elsbach 2013). For example, RottenTomatoes.com and IMDb.com agree in classifying Mel Brooks’s send-up of the western *Blazing Saddles* as <comedy> and <western>. But Rotten Tomatoes classifies *The Blues Brothers* as <action/adventure> and <comedy>, while IMDb classifies it as <comedy>, <music>, and <musical>. Our analysis considers the issues from the perspective of individual members of an audience.

## Feature Space and Label Space

The received view of classification treats labels as tied to *fixed* sets of feature values. If the relationship between labels and features remains static, then it is sufficient to model a single categorical space that individuals and organizations navigate. But once we allow the meanings of categories to change—and the possibility of change to be contingent on how actors navigate the categorical structure—it is necessary to model the potentially variable connections between the spaces. We aim to explore how actors’ movements affect the stability of social classification. So, the bedrock of our theory is a model of two spaces, *feature space* and *label space*, and the mapping that connects them. Feature space and label space are instanti-

<sup>2</sup>Scientific research provides a familiar example of producers actively playing both roles, as researchers frequently alternate between author and reviewer roles.

ations of what Gärdenfors (2004) calls conceptual spaces (see also Widdows 2004).

Research in psychology presents three standard accounts regarding the mental representation of concepts. A representation is stored as (1) a prototype, a “best example,” (2) exemplars, known images of the concept, or (3) a schema, a structured view of a concept (Smith and Medin 1981; Murphy 2002). All three perspectives agree that an object’s observable characteristics determine categorization. Research on social categories as part of a cultural system has concentrated on the third case, where concepts are stored as schemas, which tell the patterns of feature values that closely fit the concept (DiMaggio 1997). We follow this view in constructing models.

We represent producers by their positions in a feature space of observable characteristics. Producers can lie at varying distances from each other, and they move in this space when they change their feature values. In formal terms, we let  $F$  denote the space of the combinations of feature values relevant to agents in a domain.<sup>3</sup>

For this (and many other sociological applications) feature space can be represented as a weighted graph. Such a representation defines nodes as feature profiles, combinations of feature values. Suppose that there are  $K$  relevant features in a particular context. Then the set of positions is the set of an ordered  $K$ -tuple of feature values. For instance, if the features are binary, then there are  $2^K$  positions (vertices in the graph space). We can specify this type of space using a graph relation that tells which vertices are accessible from which other vertices. To represent this, the crucial issue concerns how many features can be changed at one time. Analysis is greatly simplified if we assume that only one feature value can be changed at a time.<sup>4</sup> This means that the edges in the graph connect positions that differ on only one feature value.

To fix ideas, consider Figure 1, which depicts the graph for a space defined for three binary fea-

tures (with the restriction that edges connect only those vertices that differ by one feature value). Each of the eight vertices gives a profile of feature values. For example, (0,0,0) describes the position for which each of the three relevant features has the value of zero. The edges connect vertices that differ on only one value, for example, the vertex (0,0,0) is connected to (1,0,0), (0,1,0), and (0,0,1). The vertices in this diagram are what we call *feature profiles*.

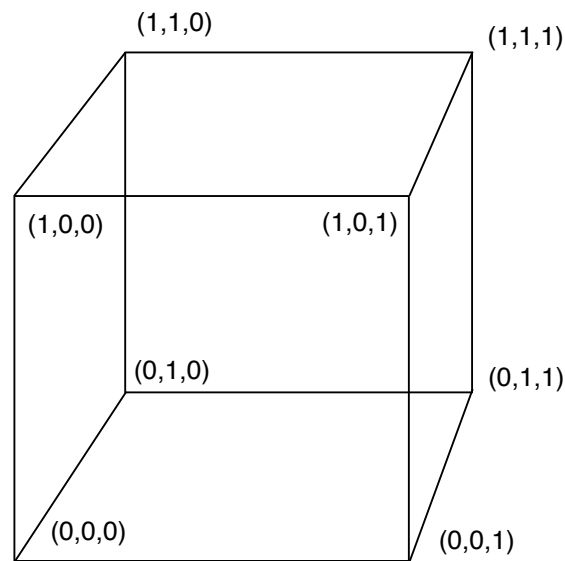
In a standard graph space, distance is given by path length, specifically the length of the shortest path that connects a pair of nodes. In our construction, each edge is associated with a transformation distance, a real-valued weight that refers to the scale of the change required to move from one configuration to another (Hahn et al. 2003). We define a distance between positions (vertices) as the scale of transformation required to convert from one set of feature values to another. So the distance from (0,0,0) to (1,0,0) is the scale of the transformation needed to change the first feature from 0 to 1. If this feature is, say, the form of authority, then a much deeper organizational transformation is required to shift from traditional to rational-legal (to use two of Weber’s types) than to change more technical features.

We treat distances as subjective judgments from the view of a focal agent and also allow the possibility that the structure of the space morphs over time. Precision demands that we tune the formal notation to fit these considerations by including argument slots for the agent and the time point in the distance function. Economy of expression argues the opposite. We use the more economical representation and omit from our notation the dependence of each predicate and function on these arguments. But each function and predicate should be considered as dependent on the agent and on the time point. With this convention we express *distance in feature space* as  $d^F(v, v')$ , the scale of the transformation required to convert from configuration  $v$  to the configuration  $v'$ . The pair  $\langle F, d^F \rangle$  defines the metric feature space.

The second space of interest is also a graph space, one whose vertices are labels. We refer to this space as a *label space*, represented as  $L$ . Distances in this space, denoted as  $d^L(l, l')$ , tell the distances between labels. Again we allow the possibility that distances between adjacent

<sup>3</sup>In our construction a feature is a dimension such as the form of authority in an organization. Positions in feature space are vectors of the values of a set of features, for example, traditional, rational-legal, and charismatic forms of authority.

<sup>4</sup>Hannan et al. (2003) discuss these issues in the context of organizational change and consider ways of relaxing this constraint.



**Figure 1:** Example of a feature space with three binary features.

vertices might differ. The pair  $\langle L, d^L \rangle$  defines the metric label space.

We need to control the scope of label space, because social actors use many kinds of labels. Some of them are related, but others are not, for example, a person might be labeled as  $\langle \text{Hungarian} \rangle$ ,  $\langle \text{logician} \rangle$ , and  $\langle \text{archer} \rangle$ . This difference matters, because movements among related categories can affect the structure in a way that moves among unrelated categories do not. For theoretical analysis, it suffices to restrict the scope of the arguments to cases in which the language has a relatively simple structure. We use the notion of a domain, a super-category whose included categories compose the set to be analyzed (Hannan et al. 2007). Within such a structure, the schemas for all of the subcategories contain the codes for the supercategory's schema (along with other differentiating codes). These categories are related within the language of the audience that makes these distinctions. The scope of our theory is restricted to a single domain.

In addition, we restrict the scope of label space to include labels at the same hierarchical level. If we do not do this, then analysis of the interdependence of movements among categories gets very complicated without shedding

additional insight. We limit the scope to sets of categories within a domain that satisfy the restriction that none of the categories considered in the analysis is a subcategory of any other. When we refer to a label, it should be considered to be the tag for a category that is an element in the common language of the focal domain.

## Relation between the Spaces

Social classification emerges from the association of labels with features through schematization. A key element of our theoretical development—one not present in other work on social categorization—is that we define explicitly the mapping between features and labels. Our conceptualization builds on the following considerations about the nature of schematization:

1. Schematization ties labels to *sets* of feature-value profiles.
2. The mapping is partial; not all feature profiles are schematized.
3. The agent's language is coherent in the sense that different labels point to distinct collections of profiles. That is, the *sets* of



profiles (not particular feature values) associated with different labels do not intersect.

4. Schematization preserves distance, meaning that distance in one space can be inferred from distance in the other.

We can make these abstract ideas somewhat more concrete with an example. Figure 2 depicts two graphs and a mapping connecting them. The feature space contains five feature profiles denoted  $a$  through  $f$  (each individual profile is a set of feature values, as illustrated in Figure 1). Again the edges connect profiles that differ on only one feature value. The spacing between vertices indicates the transformation distance between them. The label space contains three tags: 1, 2, and 3. The morphism we want to characterize maps  $a$  and  $b$  to 1 and  $c$  to 2,  $d$  and  $e$  to 3, and does not associate a label with  $f$ . The absence of edges connecting labels preserves the absence of edges between schemas. The set of schemas that have a common label are shaded in the figure. The dotted lines display the details of the mapping.

The next step is to provide a formal characterization of the relation between the spaces. Intuition suggests that the mapping be defined from feature space to label space because the existence of clusters of feature values induces labeling (Hannan et al. 2007). However, consideration of the list of desiderata (and the figure that exemplifies them) suggests otherwise. All labels are schematized, but not all combinations of feature values belong to a schema. We can obtain the first three desiderata by defining a one-to-one mapping from labels to sets of feature-value profiles. Such a mapping yields an isomorphism. If we add that the mapping preserves structure, namely, distance, then the isomorphism has the form of an embedding.

**Definition 1** (Isomorphic embedding of sets of profiles in feature space in the space of labels). Let  $\Phi \subseteq \mathcal{P}(F)$  denote a subset of the powerset of the feature-value profiles,<sup>5</sup> and  $L$  denote the set of labels used by the focal agent for the entities in the domain. We call  $\sigma$  a distance-preserving isomorphic embedding of the power set of feature profiles in the set of labels if the following

<sup>5</sup>The powerset of a set is the set,  $\mathcal{P}(\cdot)$  is defined as the set of all of its subsets. In the case of the set  $x = \{a, b\}$ ,  $\mathcal{P}(x) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$ .

conditions hold:

$$\sigma : L \mapsto \Phi,$$

and

$$\begin{aligned} &\forall \phi_1, \phi_2, \phi_3, \phi_4, l_1, l_2, l_3, l_4 [(\sigma(l_1) = \phi_1) \\ &\wedge (\sigma(l_2) = \phi_2) \wedge (\sigma(l_3) = \phi_3) \wedge (\sigma(l_4) = \phi_4) \\ &\rightarrow (d^F(\phi_1, \phi_2) = d^F(\phi_3, \phi_4)) \\ &\leftrightarrow (d^L(l_1, l_2) = d^L(l_3, l_4))]. \end{aligned}$$

The second clause ensures that the mapping preserves distance. It states, using formal language, that if the feature space distance between  $\phi_1$  and  $\phi_2$  is the same as the distance between  $\phi_3$  and  $\phi_4$ , the distance between their respective labels is also the same. This embedded isomorphism provides the following novel representation of a schema.

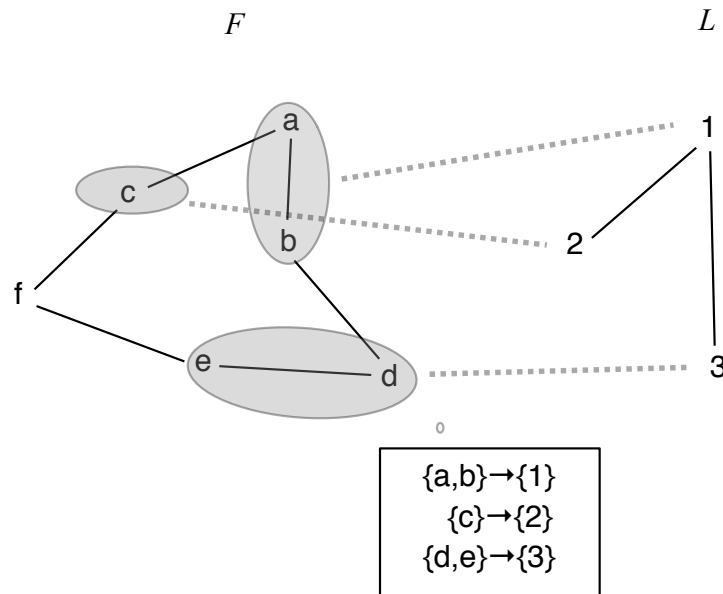
**Definition 2** (Schema). An agent's schema for a category is the subset of the feature space that the agent associates with its label (under the isomorphic embedding):  $\text{SCH}(\lambda, l, y) \longleftrightarrow \lambda = \sigma(l, y)$ .

We must add an auxiliary assumption that the agents engage in schematization along the lines of the preceding construction.

**Auxiliary assumption 1.** *The agents in the audience apply labels to schemas in a way that corresponds to the isomorphic embedding,  $\sigma$ .*

We restrict our analysis to situations in which the focal agent's schematization can be represented by a distance-preserving isomorphic embedding as in Definition 1. Instead of invoking the predicate  $\text{SCH}(\lambda, l, y)$  whenever we introduce a label in a formula, we use the notational shorthand that the symbol  $\lambda_y$  refers to a schematized label for the agent  $y$ .

A schema might be a singleton, but generally schemas have higher dimensionality. We define the distance between a point in feature space and a schema using the standard definition of the distance between a point and a set:  $d^F(f, \lambda)$  is defined as the minimum of the distances of the point to each member of the set. When it comes to the distances between schemas, we need a distance metric appropriate for sets: Hausdorff distance. In the case of the schemas  $\lambda$  and  $\lambda'$ , this metric, denoted as  $H(\lambda, \lambda')$ , is constructed as follows. First choose the element in  $\lambda$  that is



**Figure 2:** Example of an isomorphic embedding of schemas (sets of feature values) and labels.

closest to the set  $\lambda'$  and record this (minimal) distance; then calculate the minimal point-to-set distance from  $\lambda'$  to  $\lambda$ . The Hausdorff distance between the sets is the maximum of the two minimal point-to-set distances.<sup>6</sup>

It follows from our construction that the distance between a pair of labels equals the Hausdorff distance between the schemas associated with them under the isomorphic embedding. This construction yields an implicit definition of distance in label space as the Hausdorff distances between the associated schemas.

This defines an agent's schema, which links feature profiles with labels. Whether particular sets of profiles become associated with a label is another question. Of course, a theoretical model cannot predict this, because the audience determines the sets of features that become a label's schema. Recent sociological work on classification

has relied on experts, for example critics or government agencies, to assign objects to a category—or to provide the category's extension.<sup>7</sup>

Our argument does not require external agents to maintain categorical boundaries. Rather, we suggest that the link between the spaces builds on the characteristics of the objects that are already labeled (those in the category's extension). Consider the example in Figure 2. Suppose that there are three producers and that  $x_1$  has the feature profile  $a$ ,  $x_2$  has  $b$ , and  $x_3$  has  $c$ . Now suppose that a focal agent applies the tag 1 to  $x_1$ ,  $x_2$ , and  $x_3$  (so the agent's extension of this label consists of these three producers). What has happened here is that the focal agent has

<sup>7</sup>Logicians and linguists define concepts in two ways: extensional and intensional. The extension of a concept refers to the set of objects that satisfy the concept in one fixed context. The intension of a concept refers to its meaning over possibly changing contexts (alternative possible worlds). The extension of a concept is an actual set of members; in the intentional view a concept has a more abstract characterization. For instance, consider how we would define the concept <prime number>. We can give its (partial) extension as {3, 5, 7, 11, ...}. The standard meaning (or intension) of this concept is "a natural number greater than one that is divisible only by one or itself."

<sup>6</sup>More formally,

$$H(\lambda, \lambda') \equiv \max \left\{ \sup_{v \in \lambda} \inf_{v' \in \lambda'} d^F(v, v'), \sup_{v' \in \lambda'} \inf_{v \in \lambda} d^F(v, v') \right\},$$

where inf and sup refer to the infimum and supremum, respectively.



reacted to the partial membership of  $x_3$  in the agent's schema for the label and has assigned this producer membership. Other agents, on learning of the first agent's extension, might experience positive updates in the probability of associating the feature profile  $c$  with the tag 1. So the link between features and labels emerges endogenously from the positions of objects in each space. As objects change features and as label assignments change, the link between feature space and label space can also change.

## Categorization: Assigning Labels to Producers

Next we turn to the issue of categorization: the assignments of labels to objects (producers in our case). We start by considering the link between feature space and label space. We then build our model for the standard view of category statics. We subsequently extend it to consider category dynamics.

Consider the producer's location in feature space, which we denote by  $f_x$ . Refer again to the example in Figure 2. An agent will normally assign the tag 1 to a producer with feature-value profile  $f_x = a$ . But  $a$ , as drawn in this figure, lies one step from  $b$  (in the same equivalence class) and one step from  $c$  (not in the class) and the  $d^F(a, c)$  is not much greater than  $d^F(a, b)$ . Producers with the feature profile  $a$  plausibly could be assigned partial memberships in the categories tagged as 1 and 2. Likewise, the profile  $f$  (which is not associated with a label in this example) lies one step from  $c$  and one step from  $e$ —it stands between the two categories. Perhaps it will be assigned partial memberships in 2 and 3. So, a producer's distance from the schema associated with a concept influences whether it gets assigned full or partial membership.

Considering partial memberships marks a departure from classical understandings of concepts and categories. In the so-called classical perspective, categories are *crisp*: objects either bear a label or they do not, and they either fit a schema or do not.<sup>8</sup> For example, every natural number is either a prime or not, and no prime number is more prime than any other. But research in

<sup>8</sup>More precisely, in the classical view, both the extension and intention of a category are crisp.

cognitive science over the past 35 years has found that people do not classify social and material objects in such a crisp fashion. Rather, they see some objects as having partial memberships, which means that concepts can have a typicality structure. For example, apples and oranges are judged as typical <fruit> but tomatoes and olives are atypical (Rosch and Lloyd 1978). In the case of markets, an independent business that serves espresso and sandwiches might be considered “sort of” or “technically” a <restaurant> but primarily a <coffee shop>. A software producer that develops spreadsheets might have partial membership in <database>.

The notion of a graded membership or typicality refers to the degree to which the focal agent regards a producer's feature values as fitting her schema for the label (or the similarity to the prototype). Producers with feature values close to the schema/prototype are said to be typical of that agent's concept.

In models of categorization, whether a person applies a label to an object is based on the distance between her mental representation of the object and her representation of the respective concept (e.g., schema or prototype). The higher an object's typicality, the higher the probability that the agent regards it as an instance of the associated concept (Rosch and Mervis 1975; Tversky and Gati 1978). For example, restaurants with similar menus get classified alike (Kovács and Johnson 2014). This means that whether an object gets assigned a label declines monotonically with its distance from the schema/prototype. We propose a simple functional form for this relationship.

**Postulate 1.** *An object's typicality with respect to a concept is a negative logistic function of the distance of its feature profile from the agent's schema for the concept label:*<sup>9</sup>

$$\mathfrak{N}l, y [(\lambda_y \neq \emptyset) \rightarrow \exists! a, b \forall x [(a > 0) \wedge (b > 0) \wedge (Ty(l, x, y) = (1 + a e^{b d^F(f_x, \lambda_y)})^{-1})].$$

<sup>9</sup>The model we develop is inspired by Hampton (2007) and Verheyen et al. (2010) but is adapted to accommodate an argument about leniency. In particular, Verheyen et al. (2010) build a Rasch model of categorization to implement Hampton's notion of agent-level thresholds, which we do not include. The negative logistic is more suitable for our purposes because it has a classic probabilistic interpretation.

(In informal language, this postulate states that, normally, a producer's typicality to a schema-tized label  $l$  is a negative logistic function of the distance between the producer's feature profile and the audience member's schema for  $l$ . The formula specifies the existence of the positive constants included in the logistic function and that the formula holds for all labels and all audience members in the domain.)

Here and elsewhere,  $\mathfrak{N}$  denotes a nonmonotonic quantifier. Formulas quantified in this way provide formal representations of generic rules (or rules with possible exceptions). Such rules tell what is normally the case. The semantics of this quantifier are spelled out in Pólos and Hannan (2004). In addition, we use an informal sorting of variables where  $l$  refers to a label,  $x$  to an object, and  $y$  to an audience member.

## Statics

Postulate 1 has implications for labeling. First we develop the standard view, which we will call the static perspective. In this view it is assumed that both the category schema and the characteristics of objects do not change. The main question is how atypical objects are assigned membership in categories that have fuzzy boundaries.

Fuzziness in category boundaries arises because agents tend to form different extensions over occasions. Sometimes they will include an object in a category's extension and sometimes not (for objects with moderate typicality). Recognition of this tendency has led to the use of a probabilistic notion of categorization. Theory and research emphasize the probability that an object sits in the extension of a concept (over occasions). Let  $\pi(l, x, y)$  denote the probability that the agent  $y$  currently will apply the label  $l$  to the object  $x$ . We propose that this probability is proportional to  $x$ 's typicality for the concept.

**Postulate 2.** *The probability that an agent categorizes an object as a member of a category is proportional to its typicality with respect to the concept:*

$$\mathfrak{N}l, y [(\lambda_y \neq \emptyset) \rightarrow \exists! k \forall x [(k > 0) \wedge (\pi(l, x, y) = k \cdot Ty(l, x, y))].$$

This small argument implies that an object's feature-based distance from the relevant schema affects how it gets labeled.<sup>10</sup>

**Lemma 1.**

$$\mathfrak{P}l, y [(\lambda_y \neq \emptyset) \rightarrow \exists! a, b, k \forall x [(a > 0) \wedge (b > 0) \wedge (k > 0) \wedge (\pi(l, x, y) = k \cdot (1 + a e^{b d^F(f_x, \lambda_y)})^{-1})].$$

## Dynamics

Now we turn to dynamics. Here we allow that both category schemas and object characteristics can change. To model this, we need to shift from considering static probabilities of categorization to the hazards of adding and dropping labels. We turn to the question of how an object's typicality, based on its feature profile, and the changing properties of the categories affect the dynamics of labeling, the hazards of adding and dropping labels. We start by stating a straightforward extension of the typicality argument.

**Postulate 3.** *An agent's hazard of newly applying a label to a producer normally increases with the typicality of the producer's profile of feature values from the schema for the label; and her hazard of dropping the label decreases with this typicality:*

$$\mathfrak{N}l, l', y \forall x, x' [(Ty(l, x, y) > Ty(l', x', y)) \rightarrow ((l \notin \mathcal{L}_{x, y}) \wedge (l' \notin \mathcal{L}_{x', y}) \rightarrow \gamma(l, x, y) > \gamma(l', x', y)) \wedge ((l \in \mathcal{L}_{x, y}) \wedge (l' \in \mathcal{L}_{x', y}) \rightarrow \delta(l, x, y) < \delta(l', x', y))].$$

Then it follows that the hazards are determined by the distance between a feature-value profile and the applicable schema.

**Proposition 1.** *The hazard of an agent's newly applying a label to a producer presumably falls with the distance of the producer's profile of feature values from the agent's schema; and the hazard of the agent's dropping the label increases with*

<sup>10</sup>The implications of a set of rules with exceptions are the logical consequences of a stage of a theory. Such provisional theorems have a haphazard existence: what can be derived at one stage might not be derivable in a later stage. So the status of a provisional theorem differs from that of a causal story. The syntax of the language codes this difference. It introduces a "presumably" quantifier, denoted by  $\mathfrak{P}$ . Sentences (formulas) quantified by  $\mathfrak{P}$  are provisional theorems at a stage of a theory (if they follow from the premises at that stage).

this distance:<sup>11</sup>

$$\begin{aligned} \mathfrak{P} l, l', y \forall x, x' [(d^F(f_x, \lambda_l) < d^F(f_{x'}, \lambda_{l'})) \rightarrow \\ ((l \notin \ell_{x,y}) \wedge (l' \notin \ell_{x',y}) \rightarrow \\ \gamma(l, x, y) > \gamma(l', x', y)) \wedge ((l \in \ell_{x,y}) \wedge (l' \in \ell_{x',y}) \\ \rightarrow \delta(l, x, y) < \delta(l', x', y))]. \end{aligned}$$

*Proof.* The rule chain that connects the antecedent and consequent is a simple cut rule, which carries over from first-order logic to the nonmonotonic logic we use. The cut rule holds that  $\phi \rightarrow \psi$  and  $\psi \rightarrow \chi$  logically implies  $\phi \rightarrow \chi$ . Postulates 1 and 2 imply that typicality declines with distance from the applicable schema. Our extension of the static typicality argument to dynamics (postulate 3) states that higher typically increases the hazard of adopting a label and decreases the hazard of dropping a label. This completes the chain that connects the antecedent and the consequent. The argument does not support a negative rule chain, a chain whose links run from the antecedent to the *negation* of the consequent. The absence of such a rule chain means that the implication stated in the proposition is proven.  $\square$

For this proposition, we have not assumed that the schemas for labels remain fixed. Rather, an agent's schema is influenced by characteristics of the objects that currently belong to the category. This means that schemas change when members move within each space, and these changes influence subsequent categorization. Thus emerges a coupled ecology of market classification. Objects can independently change positions in feature space by changing their characteristics, or in label space by being reclassified, and this changes the link between feature space and label space. Altered category schemas then cause agents to change other categorizations, possibly leading to cascades of changes.

## Leniency and Categorization

We might also expect that label-space properties affect labeling. In contexts that lack mechanisms for boundary enforcement, labels can become

<sup>11</sup>Note that  $l$  and  $l'$  are variables—not names—and we have not stipulated that  $l \neq l'$ . So the proposition applies when there is a single label in the domain. This becomes important when we modify the relationship in meaning postulate 1 below.

porous and have vague boundaries. This might modify the mapping between the two spaces.

Categories with vague, porous boundaries proliferate in the social world. Scholars who focus on the cognitive processes underlying categorization emphasize that categories organize knowledge and enable humans to structure their reality (Murphy 2002). But sociologists who study “on-the-ground” classification have noted that loose and overlapping categories appear in all types of classification systems, even those formally constructed by authorities, what DiMaggio (1997) called administrative classification. For example, the International Classification of Diseases (ICD) retained a number of ill-defined categories for causes of death, such as < eclampsia >, < convulsion >, or < hemorrhage >, to guard against either having too many unclassified cases or inflating figures of “badly defined diseases” (Bowker and Star 2000). From a perspective that stresses the benefits of concepts in allowing economic and rapid cognition, such vague concepts might seem to be pointless and counterproductive. However, loose concepts exist precisely because classification addresses a messy reality.<sup>12</sup> Loose concepts and categories allow classification to incorporate entities that otherwise would be excluded, without jeopardizing the crispness of other concepts or categories.

We want to model how vague categories can affect the ecological dynamic introduced above. To do this, we need a formal definition of categories with vague, porous boundaries. We refer to categories that are close to many other categories as *lenient*.<sup>13</sup>

Lenient categories might not provide a basis for boundary enforcement. A key intuition is that lenient categories therefore provide less-sharp meanings. We argue that this tendency

<sup>12</sup>An extensive literature in logic, linguistics, and philosophy of science explores the role of vagueness in scientific and natural languages. See Black (1937), van Demeter (2010), and the papers collected in Keefe and Smith (1997).

<sup>13</sup>The language of leniency suggests that observed patterns of overlap reflect constraints, that the differences among categories in sharpness of boundaries and breadth of overlap are not accidental. This judgment, of course, involves an inference. In developing the core argument, we focus on the descriptive information contained in patterns of overlaps. Once the presentation of the argument is complete, we show that it implies that the patterns of overlap generally do reflect constraint (as opposed to accidental patterns of overlap).

arises from confusion caused by the similarity of lenient categories to others. For example, entrants into <disk array producer> were already identified with many other categories, including <storage subsystems>, <RAID>, and <network-attached storage>. As a result, the label took on diffuse and variegated meanings (McKendrick et al. 2003). In the emergence of the nanotechnology field, the label became quite loose due to overlaps with several scientific disciplines, including molecular physics, materials science, computer science, and electrical engineering (Grodal 2011). The possibility that a welter of categories applies to some members makes it hard for agents to interpret these memberships. We develop this notion in terms of proximities in label space.

Our argument focuses on the distance of a focal category from the other categories in the domain, which we call total distance.

**Definition 3** (Total distance of a category from all others in the domain).

$$D(l, y) = \sum_{l' \in \mathcal{L}} d^L(l, l') \\ = \sum_{l' \in \mathcal{L}} d^F(\lambda_y, \lambda_{y'}) = \sum_{l' \in \mathcal{L}} H(\lambda_y, \lambda_{y'}).$$

Lenient categories have low values of  $D$ . We do not suggest that *any* difference in leniency will be consequential. Rather, we propose that there is a threshold over which multiple overlaps make a category difficult for agents to interpret. The following meaning postulate formalizes this idea and provides an indirect definition of leniency.

**Meaning postulate 1.** *If one category is much more lenient than another, then it normally is less distant from other categories in the domain.*<sup>14</sup>

$$\mathfrak{N}l, l', y \exists w [(w > 0) \wedge (\text{len}(l, y) - \text{len}(l', y) > w) \rightarrow D(l, y) < D(l', y)].$$

The relationship just postulated can hold for many values of  $w$ . We identify (and name as  $w^*$ ) the minimum value for which the relationship holds:

$$w^* = \inf\{w \mid (w > 0) \wedge \mathfrak{N}l, l', y [(\text{len}(l, y) - \text{len}(l', y) > w) \rightarrow D(l, y) < D(l', y)]\}.$$

<sup>14</sup>In the nonmonotonic logic we use, postulates can be strict (that is, first-order rules without exceptions) or generic statements that hold in normal situations, rules with possible exceptions—see the supplementary appendix.

We suggest that leniency affects the link between feature-space positions and labeling. This is because categories do not exist in isolation. The vast research literature on concepts builds on the notion that categories play the cognitive role of grouping similar items and differentiating different items. Categorization concerns not only the match between an object and an agent's schema for one category but also the difference of the object from the agent's schemas for other categories. A metaphor can be seen in the concept of *cue validity*, which characterizes the diagnosticity of a feature value for membership in a particular category. Cue validity is based on both how much a feature is associated with a focal category and how little it is associated with *others* (Rosch and Lloyd 1978). Analogously, an object's fit to a schema is more diagnostic of membership if the schemas for other categories are not proximate.

Central to our argument is that having high proximity to many categories distorts perceptions of categorization based on feature values. To develop this formally, we first define, for producer-category pairs, the feature-space proximity between a focal producer and other producers assigned labels of all categories *other* than a focal category.

**Definition 4** (Overall proximity to other-labeled objects at the producer level).

$$P(l, x, y) = [\sum_{l' \neq l} \sum_{l' \in \ell(x', y)} d^F(f_x, f_{x'})]^{-1}.$$

We then need to modify the standard assumption about how agents categorize (see Postulate 1). Specifically, we propose that high overall proximity weakens an object's fit to categorical schemas.<sup>15</sup>

**Postulate 4.** *When the domain contains multiple labels, the typicality of a producer is a decreasing function of its proximity to the members of other categories in the domain and a decreasing function of its distance from the schema for the focal category:*

$$\mathfrak{N}l, y [(N_{\mathcal{L}, y} > 1) \wedge (\lambda_l \neq \emptyset) \rightarrow \\ \exists! \alpha, \beta \forall x [(\alpha > 0) \wedge (\beta > 0) \wedge (Ty(l, x, y) \\ = (1 + e^{\alpha P(l, x, y)} e^{\beta d^F(f_x, \lambda_l)})^{-1})],$$

<sup>15</sup>In introducing this consideration, we introduce a specificity consideration: that the postulate that follows holds only when the domain contains more than one label. In the following we use specificity considerations to manage the inconsistency between the two indirect definitions of typicality.



where  $N_{\mathcal{L},y}$  denotes the number of domain labels that the agent has schematized.

The typicality function in postulate 4 approaches the function given in postulate 1 as  $P \downarrow 0$ . As  $P$  increases, typicality falls toward zero. This means that the higher an object's proximity to other categories, the weaker is the effect of its distance from the schema on typicality. Figure 3 illustrates the effect of increasing proximity in terms of this function. The upper (solid) curve plots  $\mu = (1 + 0.02e^{2d})^{-1}$ , and the lower (dashed) changes the value of the multiplier of the exponential term from 0.02 to 0.05, reflecting a much higher value of overall proximity. For the solid curve, where the producer has low proximity to members of other categories, an object at close distance to the schema results in a high expected typicality. For the dashed curve, where the producer has high proximity to members of other categories, an object at the same distance from the schema will have a lower expected typicality.

While the co-presence of postulates 1 and 4 in the theory would create an inconsistency in first-order logic, this need not be so for nonmonotonic logic. We resolve this potential conflict by relying on the most characteristic feature of nonmonotonic logic that more specific rules override less specific ones. By restricting the postulate to cases in which a domain contains more than one label, we have made postulate 4 more specific than postulate 1. This more specific relationship overrules the less specific one in the analysis of multilabel domains.

Next we tie the argument about labeling to leniency. The first step is to define producer's closest category. It simplifies the analysis greatly, without distorting the intuition, to assume that the closest category is unique for each producer.

**Definition 5** (Closest category).

$$\text{CL}(l, x, y) \leftrightarrow \forall l' [(l' \in \mathcal{L}) \rightarrow d(f_x, \lambda'_y) > d(f_x, \lambda_y)].$$

Next we need to tie producer-level arguments to the category-level ones. We do so by considering pairs of producers that are alike in one important respect: the distance of their feature-value profile to the schema of their closest categories is the same. In other words, we consider pairs who

fit the most applicable (best-fitting) schema to the same degree. But the pairs also differ in an important respect: their closest categories differ in total distance to *other* categories.

**Postulate 5.** *Producers that are proximate to the schemas of categories that lie closer to other categories in the domain normally have higher total proximity to the members of all other categories in the domain (holding constant the degree of fit to the schema for the closest category):*

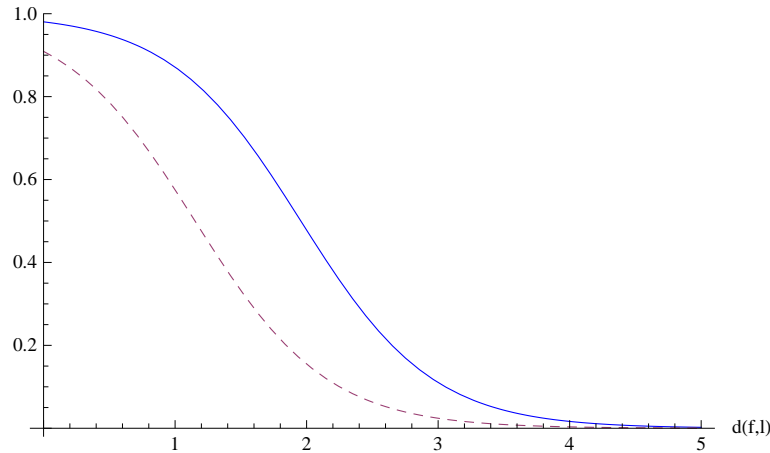
$$\begin{aligned} \mathfrak{N}l, l', y \forall x, x' [(D(l, y) < D(l', y) \wedge \text{CL}(l, x, y) \\ \wedge \text{CL}(l', x', y) \wedge (d(f_x, \lambda_y) = d(f_{x'}, \lambda'_y)) \\ \rightarrow P(l, x, y) > P(l', x', y)]. \end{aligned}$$

Together, postulates 4 and 5 imply that producers that are close in feature space to a lenient label are seen as less typical of that label, as compared to producers who are equally proximate to a constraining one.

**Lemma 2.** *If two producers are equally close to the schemas for their closest categories and the closest category for one is sufficiently more lenient than the closest category for the other, then the producer with the more lenient closest category presumably has lower typicality in its closest category than does the other producer:*

$$\begin{aligned} \mathfrak{P}l, l', y \forall x, x' [(len(l, y) - len(l', y) > w^*) \\ \wedge \text{CL}(l, x, y) \wedge \text{CL}(l', x', y) \wedge (d(f_x, \lambda_y) = \\ d(f_{x'}, \lambda'_y)) \rightarrow Ty(l, x, y) < Ty(l', x', y)]. \end{aligned}$$

*Proof.* The antecedent requires that the domain contains more than one label because it refers to a pair of labels that are applied to a pair of producers such that each bears one label but not the other. This means that specificity considerations dictate that the *proximity-weighted* typicality function (from postulate 4) applies. Given that the antecedent specifies that the producers' feature profiles stand at equal distance from the applicable schemas, any differences in typicalities depends only on the difference in total proximities for the producers. The antecedent also tells that the difference in the leniencies of the applied categories exceeds  $w^*$ , which means that postulate 5 implies that the producer close to the lenient label has higher total proximity to the members of other categories. Inserting this inequality in proximities into the proximity-weighted typicality function completes the chain



**Figure 3:** Two illustrative membership functions showing the effect of a difference in proximity on the relationship between distance from a category schema/prototype and the probability of labeling; see text.

linking the antecedent and consequent. In the absence of any opposing rule change given the premise set, this completes the proof.  $\square$

It follows that the link between a producer's proximity to a label and categorization is stronger for constraining labels and weaker for lenient labels.

**Proposition 2.** *Agents have higher hazards of adding and lower hazards of (newly) dropping the label of a producer's closest category when it is constraining as compared to when it is sufficiently more lenient:*

$$\begin{aligned} & \mathfrak{P} l, l', y \forall x, x' [(len(l, y) - len(l', y) > w^*) \\ & \quad \wedge CL(l, x, y) \wedge CL(l', x', y) \wedge (d^F(f_x, l) \\ & = d^F(f_{x'}, l')) \wedge ((l \notin \mathcal{L}(x)) \wedge (l' \notin \mathcal{L}(x', y))) \\ & \rightarrow \gamma(l, x, y) < \gamma(l', x', y)) \wedge ((l \in \mathcal{L}(x, y)) \\ & \wedge (l' \in \mathcal{L}(x', y)) \rightarrow \delta(l, x, y) > \delta(l', x', y))]. \end{aligned}$$

*Proof.* The rule chain supporting lemma 2 entails a difference in typicalities for the pair of producers. A chain rule with the extension of the standard typicality argument to the dynamic situation, postulate 3, provides a positive rule chain linking the antecedent and consequent. The theory as stated does not give rise to any negative rule chains that imply a different conclusion (a different ordering of the hazards of dropping labels). Thus the nonmonotonic proof is complete.  $\square$

The argument that supports proposition 2 holds that a category's structural position in

label space affects how strongly it is linked to feature space. Specifically, the relationship between feature-space positions and labeling weakens when labels are lenient. This provides a theoretical basis for predicting loose coupling of feature values and label assignments.

Perhaps the most interesting implication of this argument concerns the trajectories of leniency and constraint in classification. Our argument suggests that lenient categories become more lenient: agents are less likely to take into account feature-based similarities when applying lenient labels, which can lead to a diverse set of agents being assigned a lenient label. This will increase the proximity of its members to the memberships of other categories, which will fuel the cycle. This pattern does not hold for members of much more constraining categories. The “Discussion” section sketches the argument (for a fixed set of categories) that lenient categories remain lenient and highly constraining categories remain constraining.<sup>16</sup>

## A Context: Market Classification in the Software Industry

We investigate this model in the empirical setting of the software industry. In the software

<sup>16</sup>We add the qualification to fixed sets of categories because new categories might deviate from the pattern.



industry, an informal classification of product markets has emerged through interaction among producers, analysts, the media, and consumers. It depends heavily on producer affiliation with market categories. Consumers, employees, and investors use category labels to find producers, products, and services, as well as to stay abreast of innovations. Organizations identify with labels primarily to communicate what they offer to various audiences.

In making the shift to a specific context, we also shift from a single agent to a collective, an audience. This is because we relate labeling actions by a focal firm to similar decisions by others. These analyses depend on an assumption that the language is common within the audience, otherwise it would be feckless to search for the kinds of effects our theory yields. The degree to which this assumption characterizes audiences of interest is an interesting and important topic in its own right (Pontikes 2012; Koçak et al. 2014). But we do not address it here. What is important is that the audience investigated, software producers, uses a common system of classification to make sense of the domain.

We represent label space based on affiliations with market categories made by software producers in press releases. Press releases are an important medium for an organization to convey what it does. Software organizations use press releases to distribute news and create a public face for media analysts, investors, and consumers. The major stock exchanges in the United States suggest or require that companies issue press releases to distribute information. Media often note press releases in their coverage: a study of public companies finds an average of 1.5 (median 1.3) media articles per press release issued from 2001 to 2006 (Soltes 2009). Product classification is based on attributes of a company's product or service, so labels in this context are verifiable by consumers or analysts. As a result, it is important for organizations to describe their activities accurately. Because press releases can be produced at low cost, even small and young organizations (which are difficult to track otherwise) issue them. For these reasons, press releases provide a good source of data to construct label space in this domain.

Software producers use category labels to describe themselves in press releases. For exam-

ple, MicroStrategy identified itself in 1995 as "a leading supplier of client/server decision support tools," thereby affiliating with two category labels. The <client/server> label indicates that MicroStrategy's software would run on multiple client machines that communicated through a server. The <decision support> label conveys that the software provides reports that combine data from a number of sources to identify and solve problems and support business decisions (Figure 4 lists other examples of label affiliations from press releases in these data). In this context, producers typically develop a suite of software products targeted at one or more market categories, and the category affiliation is at the level of the organization.

We investigated whether labels claimed in press releases were also used by other important audiences. We found that the preeminent industry-analyst organization, Gartner, issued reports on over 50 percent of labels extracted from press releases, showing that press releases capture a common language for classification among audiences in this domain. Of the labels used in both press releases and in Gartner reports, over 75 percent appeared first in press releases.

We also investigated whether affiliations in press releases target a specific audience or whether they reflect the general identity of the organization. This question is especially important because different audiences can have different preferences regarding classification (Pontikes 2012). We compared claims claimed in press releases and company websites from the same time period (recorded in the archived web) and those claimed in annual reports (10-K forms) for public companies. Organizations identified with the same labels in press releases and on websites 71 percent of the time and in 10-K forms 81 percent of the time. This indicates that label assignments from press releases reflect an organization's general market identity.

This classification contains both lenient and constraining categories. Constraining categories evoke specific expectations. For example, <entertainment software> refers to producers of video game software. This category has a trade association, the Entertainment Software Association (or ESA), that tracks membership and promotes the interests of members, such as lobbying the government on issues like video game ratings. This

Organization	Date	Description
Watson General	May 1994	Watson General currently provides remote software monitoring systems.
American Software	May 1994	American Software develops, markets and supports the industry's most comprehensive offering of integrated supply chain management systems.
MicroStrategy	January 1996	MicroStrategy is the leading provider of relational OLAP (ROLAP) products and services for developing and accessing enterprise data warehouses.
VCON	October 1996	VCON is one of the leading manufacturers and marketers of desktop videoconferencing hardware and software products in the industry.
TSSI	May 1999	TSSI is a leading provider of test automation software technology and solutions.
Cognos	February 1999	Cognos Corporation, the leading developer of enterprise business intelligence and application development tools.
Accrue Software	October 1999	Accrue Software, a leading provider of e-business analysis software and services.
Citrix Systems	February 2000	Citrix Systems, Inc. is a global leader in application server software and services.
Axiom	April 2000	Axiom Corporation is a global leader in real-time customer data integration and customer relationship management.
Plasmon	August 2000	Plasmon, a leading manufacturer of automated data storage solutions, today announced its Diamond® storage management software.
Ilhey software	September 2000	Ilhey Software Inc., a leading provider of contact center and voice technologies
CardSoft	November 2000	CardSoft Inc., a leading provider of open platform software on small electronic devices.
Veridicom	November 2001	Veridicom, Inc. is a leader in fingerprint-based biometrics solutions.

**Figure 4:** Sample press release identity statements with category affiliations.

category has very low overlap with others. Companies that produce other types of software for entertainment—such as digital video or audio—do not affiliate with this label. Lenient categories broadly overlap with many other categories, and their meaning is a frequent source of discussion among industry insiders. For example, industry analyst James Kobiellus of Forrester Research describes in a blog post titled “What’s Not BI?” how audience members struggle to struggle to derive meaning from the lenient <business intelligence (BI)> category:

One of the favorite pastimes of BI analysts everywhere . . . is defining and redefining this uber-category known as BI. What exactly is it? Or rather—considering that almost every data

management technology has been swept into BI’s gravitational orbit at one time or another by somebody somewhere—what is BI not? What’s analytics? What are decision support systems? (Kobiellus 2010)

Lenient labels are vague and have porous boundaries, but they are not necessarily marginal or declining in importance. In many cases, they grow large and become prominent. For example, the lenient <customer relationship management> label received 149 mentions in the *Wall Street Journal* from its first appearance in 1998 through 2008. But <electronic design automation>, a prominent constraining label, was mentioned only 33 times from its first appearance in 1988 through

2002 (its last mention). Some of the most important categories in this classification are lenient.

Our model requires that label space only include labels at the same hierarchical level. Labels analyzed should not be supercategories or subcategories of one another. This condition holds for the labels included in this analysis. Lenient categories included in this analysis are not superordinate categories. In a hierarchy, a superordinate category sits above all subcategories in a tree, such that if an organization belongs to a subcategory, it also belongs to the supercategory. This type of hierarchical relationship is uncommon in our empirical context, as is often the case for informal classification. For example, the extension of <business intelligence> overlaps with that of <sales-force automation>, but not all <sales force automation> members are labeled as <business intelligence>. Furthermore, some organizations only affiliate with lenient categories. Lenient categories overlap broadly, sometimes with more than 100 other categories; but the categories they overlap with are not nested.

## Data and Methods

### Label Space

Testing these propositions requires data on label assignments. We use data compiled by Pontikes (2008) based on identity statements in press releases issued by software organizations (for examples, see Figure 4). This identifies times at which organizations adopt new labels and drop existing labels.

A compilation of press releases issued between 1990 and 2002 in *Businesswire*, *PR Newswire*, and *Computerwire* with at least three mentions of the word “software” served as the initial source of data: 268,963 press releases. A combination of custom-coded programs for text matching and visual examination of the outputs of these programs yielded records for 4,566 software organizations that issued press releases during this period. The identity statements made by these organizations indicate category membership. Pontikes (2008) assembled an extensive list of software labels from articles in *Software Magazine* and *Computerworld* and from inspection of the firms’ identity statements, and used text-matching programs to search all identity statements for these

labels. This created a file of organizations’ labels for each year during 1990–2002. The final data contain information on 456 labels and 4,566 organizations over 18,192 organization-years.

### Feature Space

Our empirical analysis positions organizations in a feature space based on technical capabilities, which we call *knowledge space*. This space is a subset of all feature values that characterize organizations. We use patents to represent knowledge space. Patents do not capture the complete set of relevant features in this domain, but patented technology reflects an important subset of innovative activity among software producers. Patents have been widely used in previous research to identify an organization’s technological position and to measure relative similarities in a technical space (Jaffe 1986 1989; Podolny et al. 1996; Stuart and Podolny 1996; Sørensen and Stuart 2000).

Technological innovation has been at the heart of the software industry’s growth, and a fair amount of inventive activity can be traced through patents. Patents influence venture-capital financing, IPOs, and successful acquisition (Mann and Sager 2007; Cockburn and MacGarvie 2009). Even small organizations patent, hoping to exclude competitors (Mann 2005). Despite controversy about the relevance of patents in the software industry, a sizable number of software organizations received patents (20 percent in the press release data).<sup>17</sup>

Patents provide a good basis to measure positions in a technological feature space, for several reasons. First, they reflect an organization’s technical capabilities. Second, a independent party verifies them. Third, the patenting process is very weakly influenced (if at all) by an organi-

<sup>17</sup>Patents have been controversial in software. *Gottschalk v. Benson* in 1972 ruled that software programs were algorithms that could not be patented. However, this decision was mostly overturned in *Diamond v. Diehr* in 1981, which found that a software program could be patented if it was embedded within an apparatus. After a series of cases that increasingly supported the patentability of software, a 1995 ruling essentially reversed *Gottschalk*, and the last barrier to patenting pure software was overturned in 1998. Despite this, the approval of software patents was a routine practice long before the courts recognized it (Cohen and Lemley 2001; Cockburn and MacGarvie 2009; Mann 2005).

zation's product-market categorization. In this context, market categories are only loosely based on technologies. For example, <collaboration software> combines technologies based on video, GUI (graphical user interface), and mathematical algorithms. Feature-space data based on product evaluations by industry analysts would be much more likely to be influenced by how the organization positioned itself using product-market classification (not to mention that these data are not systematically available). A knowledge space based on patents captures a critical subset of feature space for this domain that can be measured independently from label space. Through citations, patents provide a historical record of knowledge-space positions and proximities.

The downside to using patents is that not all software organizations patent. We conduct our main analyses on the subset of organizations that are active in knowledge space (those that have previously patented). The press release data contain 789 organizations with at least one patent across 4,012 organization-years. We run additional analyses on all organizations in the press release data, coding those that are not active in knowledge space (those without patents) as having zero knowledge-space proximity to all labels. The results are consistent across both analyses.

Patent citations come from the National Bureau of Economic Research patent-data project (Hall et al. 2001). Data were updated to include patents through 2002.<sup>18</sup> The U.S. Patent and Trademark Office issues patents for new, unique, and nonobvious inventions. All patents must cite the relevant "prior art" on which their invention builds, and the prior-art citations indicate the knowledge foundation for the patent at hand. Two patents that cite the same patent as prior art are more similar in knowledge space than pairs that lack common citations.

The patent office requires that inventors' claims be focused and narrow. Inventors must cite any relevant patents of which they are aware, and the patent examiner can also add citations to ensure comprehensive citation to prior art (Alcácer and Gittleman 2006). For some studies, it is important that the citations to prior art accurately

reflect the inventor's knowledge; and citations added by the examiner are problematic. This does not pose a problem for our study, because we use patents to locate organizations in a knowledge space. Citations added by an examiner help refine the patents' positions.

Patent officers assign patents to a main class and one or more subclasses. There are about 400 classes. The National Bureau of Economic Research has created a higher-level classification system of six classes: Chemical, Computers and Communications, Drugs and Medical, Electrical and Electronics, and Other (Hall et al. 2001). This study uses all patents granted in the Computers and Communications class between the years 1990 and 2002 to construct knowledge space for the software industry. This broad class includes patents relevant to software.

Following Pontikes (2008), we construct knowledge space using a five-year window of all patents in the Computers and Communications class. We use all patents relevant to software, including those issued to software organizations, individual inventors, universities, and nonsoftware organizations over a moving five-year window, including patents that were applied for (that were subsequently granted) in the current year and four years prior.

## Independent Variables

Here we provide the details about the measurement of the two key explanatory variables.

*Leniency.* To represent leniency, we need to measure the proximities among categories in the domain. We adopt Pontikes's (2008) measure, which uses category overlaps in label space to indicate similarity between categories. A category is more similar to another in label space if many of its members also belong to the other category.

Leniency is the product of contrast and a positive function of the number of categories with which the focal category overlaps. Contrast is the average label-based typicality of its members. It measures the extent to which a label comprises typical or atypical members (Hannan et al. 2007).<sup>19</sup> To measure contrast, we first must compute label-based typicalities,  $\vartheta(l, x, y)$  from label

<sup>18</sup>The data updates can be found at <http://elsa.berkeley.edu/~bhall/patents.html>. Data were first updated through 2002, then through 2006. This research uses the 2002 update.

<sup>19</sup>It is important to incorporate contrast into the measure of leniency. Otherwise the membership of a few members would have outsized influence. For example, a



assignments. We assume that an object that is assigned to multiple categories is less typical of each and that this diminution is exacerbated when the categories lie far apart in the label space. For example, an establishment that is both a French restaurant and a car wash would be very atypical of each label.

In accordance with our theory, we take into account producers' *partial* affiliations with labels. What we call label-profile typicality, denoted by  $\vartheta(l, x, y)$ , is based on (1) the number of times a producer affiliates with a focal category in a given time period, compared to the number of times it affiliates with other categories, and (2) whether a producer is a member of multiple similar or multiple distant categories. The more a producer affiliates with a focal category, the higher its label-profile typicality. The less it affiliates with other categories—especially other *dissimilar* categories, the higher is its typicality in the focal category. Label-profile typicality ranges over  $[0, 1]$ . This construction follows Kovács and Hannan (2011) but uses a generalization that builds on Pontikes (2008). The first step calculates the similarities of all pairs of categories, in notation  $\text{sim}^L(l, l', u)$ , based on how many organizations are assigned membership in both categories in the time interval  $u$ . The second step uses a Shepard (1987) construction to calculate distance between a pair of categories based on label profiles, in notation,  $\tilde{d}_L(l, l', u)$ , as a Euclidean distance, as follows:

$$\forall y, u \exists h [(h > 0) \wedge \forall l, l', \\ t[(\text{sim}^L(l, l', u) = e^{-h \tilde{d}_L(l, l', u)})]. \quad (1)$$

We use measured similarities (based on co-occurrence) and a setting for the parameter  $h$  to obtain  $\tilde{d}_L$ . After experimentation with values for  $h$  ranging over the integers from 1 to 5, we found that similar patterns of effects of theoretically relevant variables (significance levels do vary, however). We set  $h = 3$ , which provides the best overall model fit.

The next step calculates typicality,  $\vartheta(l, x, u)$ , from label profiles using the idea that a producer's typicality in applied categories declines with the application of each new label, and the decline is greater when the newly added label lies far from the already applied categories. We calculate label

label with a hundred typical members and one member with broad overlap would appear as lenient.

profiles by taking into account that firms often release multiple press releases at multiple times during a year, and each provides information on labeling. Rather than only including *whether* a firm claims a label in a given time period, we weight the relative frequency of claims to each label. For example, if an organization claims label  $A$  in one press release and  $B$  in three press releases in the same year, for a total number of 4 label mentions, then we calculate  $l(A, x) = 1/4$  and  $l(B, x) = 3/4$ . We calculate  $\vartheta$  as follows:<sup>20</sup>

$$\vartheta(l, x, u) = \frac{\ell(l, x, u)}{\ell(l, x, u) + \sum_{l' \in \mathcal{L}} \ell(l', x, u) \cdot \tilde{d}_L(l, l', u)}. \quad (2)$$

A label's contrast is the average label-based typicality of its members:<sup>21</sup>

$$C(l, u) = \frac{\sum_{x: \vartheta(l, x, u) > 0} \vartheta(l, x, u)}{|\{x : \vartheta(l, x, u) > 0\}|}. \quad (3)$$

We calculate leniency as the product of inverse contrast and a positive function of the number of categories overlapped:

$$\text{len}(l, u) = (1 - C(l, u)) \cdot \psi(n_{l, u}) \\ = (1 - C(l, u)) \cdot \ln(\sum_{l'} \tilde{d}_L(l, l', u)), \quad (4)$$

where  $n_{l, u}$  denotes the number of other categories whose memberships overlap with  $l$ . The function  $\psi(n_{l, u})$  is computed based on a summation of the distance between  $l$  and each label  $l'$  with which it overlaps. We use a log transform because the interpretability of a label ought to drop more sharply when the number of overlaps rises in the lower range than at higher levels. Our decision to weight the number of categories overlapped by their distances from the focal label follows the spirit of the construction of  $\vartheta$ , namely, that combining dissimilar categories causes a greater loss in distinctiveness than combining similar ones.

We separate leniency into two pieces, instead of using the continuous measure, because the key propositions depend on one label being *sufficiently* more lenient than another. We break the distribution of leniency at 2.7, the approximate mid-point of the leniency of producers' labels,

<sup>20</sup>Kovács and Hannan (2011) treat  $\ell(l, x, u)$  as a dummy variable; Pontikes (2008) implicitly assumes all  $\tilde{d}_L(l, l', u) = 1$ .

<sup>21</sup>We use the standard notation  $|\cdot|$  to denote the cardinality of a set, the number of distinct elements.

and treat high (low) leniency as values above (below) the breakpoint.

*Producer's proximity to labels in knowledge space.* The other main variable to be defined is the proximity (inverse distance) of a producer's position in knowledge space from a label. Recall that we assume that the hazards of adding and dropping labels are negative exponential functions of the distance of an object's position in feature space from the schema for the label (Postulate 4). We lack knowledge of schemas. So we proceed indirectly by measuring a label's position in feature space in terms of the set of positions of the objects that lie in the label's extension. Then we measure an object's position in feature space relative to the label's extension.

We measure positions in knowledge space using their patents' prior art citations. Patents that share citations build on similar knowledge. Therefore we use citation overlaps to position patents relative to one another in knowledge space (Podolny et al. 1996). We use the patent citation network of all patents in the Computers and Communications class (Hall et al. 2001). We measure similarity from prior-art citations (Podolny et al. 1996; Podolny and Stuart 1995). The similarity of patent  $i$  to patent  $j$ ,  $\alpha_{ij}$ , is defined as  $s_{ij}/s_i$ , where  $s_{ij}$  denotes the number of shared citations between patents  $i$  and  $j$ , and  $s_i$  denotes the total number of citations by patent  $i$ .<sup>22</sup> We can capture more fine-grained detail by also considering

<sup>22</sup>We have the choice of whether to use asymmetric or symmetric measures of similarity. With asymmetric similarities, patent  $i$  can be more similar to patent  $j$  than  $j$  is to  $i$ . Some research in cognitive psychology supports the use of asymmetric similarities. Studies show that people's assessments of similarity is asymmetric across a number of domains (countries, figures, letters, signals) (Tversky 1977). Competition can also be asymmetric, for example, a specialist competes more strongly with a generalist than vice versa; because of this asymmetric measures have previously been used to define technological niches among firms using patent citation data (Stuart and Podolny 1996). There are also reasons to use symmetric similarities (defined as the number of overlapping citations divided by the number of citations made by both  $i$  and  $j$ :  $\alpha_{ij} = s_{ij}/s_i s_j$ ). Symmetric similarities translate into symmetric distances in feature space, which means that distances between two points in feature space will not depend on directionality. This is an important property for our formal model of feature space and label space. The choice does not materially affect our results. Weighing these options, we chose to use asymmetric patent similarities in our primary analysis to comport with standards of existing empirical research.

second-degree similarity, which takes into account whether a patent stands within two degrees of separation from another patent. This brings into the picture common dependence on relevant patents that were issued to nonsoftware organizations such as universities. We compute second-degree similarity between of patent  $i$  to  $k$  by multiplying their similarities to all third patents,  $j$ , and choosing the  $j$  that maximizes the similarity of  $i$  and  $k$ :  $\rho_{ik} = \max_{j|j \neq i, j \neq k} (\alpha_{ij} \cdot \alpha_{jk})$ .

The degree to which interested agents regard a patent as tied to a label presumably depends on how strongly the organization that issued the patent is affiliated with the label. If the inventor firm has a low level of typicality in a label, its patents should similarly be considered only partially representative of the label. To reflect this view, we define a patent's typicality in a label as equal to the label-based typicality in  $l$  of the firm to which the patent was issued, in notation  $\psi(l, i, u) = \vartheta(l, x, u) I(i, x, u)$ , where  $I(i, x, u)$  is an indicator variable equal to one if patent  $i$  is a member of the set of organization  $x$ 's patents (in notation  $\mathbf{p}_{xu}$ ) and equal to zero otherwise. The proximity (inverse distance) of an organization's position in feature space to a label,  $p(l, x, u)$ , is

$$\frac{\sum_{i \in \mathbf{p}_{xu}} \sum_{j | (j \notin \mathbf{p}_{xu})} \rho_{ij} \cdot \psi(l, j, u)}{|\mathbf{p}_{xu}|}, \quad (5)$$

where the second summation is restricted to the patents affiliated with the label  $l$ . We also use Equation 5 to calculate the proximity (inverse distance) of a producer's position to all labels that it currently claims,  $p^C(x, u)$ , and to all labels it currently does not claim,  $p^{NC}(x, u)$  (to be used as control variables). Because of skew in the distribution of the distances, we use (natural) logarithmic transformations of these variables in our empirical analysis.<sup>23</sup>

<sup>23</sup>The distance between an organization's patents and the patents associated with a category is a distance between two sets. As such, it might seem that we should use the Hausdorff distance, which we employ in our formal theory for distances between schemas. However, in this case, an organization can belong to multiple categories, so its set of patents can include those that are not relevant to category  $i$ , but are relevant to a category  $j$ . In this case, the Hausdorff distance would only consider the patent that was farthest from the set of patents associated with a category, which often will refer to a patent that is relevant to a different category.



**Table 1:** Descriptive Statistics for Tests Involving the Hazard of Adding a Label. ( $N = 1,434,096$ )

Variable	Mean	Std. Dev.	Min.	Max.
Org. adds a label	0.004	0.060	0	1
K-space prox. to focal label	0.006	0.045	0	2.543
K-space prox. to const. focal label	0.004	0.038	0	2.543
K-space prox. to len. focal label	0.002	0.024	0	2.073
K-space prox. to other labels org. not claimed	0.745	0.934	0	4.003
K-space prox. to labels org. claims	0.105	0.251	0	2.462

A producer's distance from a label's feature profile can change in three ways: (1) it can move away from a stationary profile, (2) other producers can move while the focal producer remains in place, or (3) both the focal producer and the other profile members move but in different directions. To explore this in supplementary analyses, we also create a measure of an organization's stationarity in knowledge space based on citation overlap between the organization's patents in the current year and its patents in the previous four years.

## Controls

Social factors, such as whether a category is resource-rich or otherwise popular, can also affect label adoption. People and organizations sometimes engage in herding behavior, where high levels of adoption in one period lead to increased rates of adoption in the next period (Strang and Soule 1998). This effect can result either because agents follow each other's actions, as in the case of fads (Lieberman 2000; Meyersohn and Katz 1957), or because the number of organizations in a category signals an unmeasured factor, for example demand for a specific product. To account for these effects, we include the label's fuzzy density (the number of organizations that are members of the label weighted by typicality) to control for the label's size in the previous period. The rate at which a label is catching on can also affect adoption patterns (Berger and Le Mens 2009). Therefore we also control for the number of new affiliations with the label and number of drops in the previous period, both weighted by typicality. We also include controls for the leniency of the label and tenure of the label measured since the beginning of our records, 1990.

Organization-level controls include the number of labels in which it has nonzero membership, the number of other labels in the industry, the count of its patents (to account for general inventive activity), and the time since the organization has last added or dropped any label. Venture capitalist investors may influence an organization to add or drop market categories (Pontikes 2012). Therefore we also control for whether the organization has recently received venture capital financing. Because old and large organizations might not pay as much attention to feature-based fit, perhaps because they receive less scrutiny, we control for the organization's tenure in our data, to account for age, and for whether it was ranked in *Software Magazine's* top 500 software companies (based on revenue) to account for size. To test whether there is a systematic relationship with knowledge-space proximity to a label, we also estimated specifications that include interactions between these variables and proximity.<sup>24</sup>

Tables 1 and 2 provide descriptive statistics for the independent variables in our analysis.

## Dependent Variables

We analyze events of adding and dropping labels in continuous time. We code adding a label as occurring in the first year in which an organization lists that label in press releases, and the dropping of a label as occurring in the last year an organization lists it (if the observation is not right-censored at that time).<sup>25</sup> Seventy percent

<sup>24</sup>These interaction effects are not significant, indicating that there is not a detectable difference in the influence of feature similarities on label assignments based on organization age or size. Importantly, the key effects persist.

<sup>25</sup>Label claims from the initial year an organization appears in press releases are not counted as an add, and claims from the final year are not counted as a drop.

**Table 2:** Descriptive Statistics for Tests Involving the Hazard of Dropping a Label. ( $N = 10,908$ )

Variable	Mean	Std. Dev.	Min.	Max.
Org. drops a label	0.403	0.491	0	1
K-space prox. to focal label	0.047	0.144	0	2.499
K-space prox. to const. focal label	0.024	0.106	0	2.499
K-space prox. to len. focal label	0.022	0.104	0	1.818
K-space prox. to labels org. is not claimed	0.915	0.965	0	4.003
K-space prox. to other labels org. claims	0.731	0.928	0	4.014

of organizations add and drop at least one label during the time period studied. Ninety-five percent of patenting organizations add and drop at least one label. Twenty percent of all organizations and 45 percent of patenting organizations add more than five labels. Twenty percent of all organizations and 40 percent of patenting organizations drop more than five labels.

### Stochastic Specifications and Estimation

We test these hypotheses using event-history analysis. We estimate the effects of proximity and leniency on the hazards with standard piecewise-constant specifications. We updated covariates for each time piece. In both forms of analysis, pieces are defined for less than one year, [1–2) years, [2–4) years, and 4 years or greater. Standard errors are clustered by label.

*Adding labels.* In analyses of the hazard of adding a label, the risk set consists of all dyads of an organization that has previously patented and one of the labels with which it does *not* affiliate. There are 337,800 organization-label dyads over 1,434,096 organization-label-years, and 5,266 events of adding a label. We also conduct additional analyses using data including all organizations in the press release data, whether they patent or not. In this case, we assign nonpatenters knowledge-proximity scores of zero for all labels. This larger data set contains 1,893,569 organization-label dyads over 6,485,521 organization-label-years and 15,103 events of adding a label. Dyads enter the model in the first year the organization or label enters the data (whichever comes later). Duration is the time since the organization-label dyad exists in these data.

*Dropping labels.* In analyses of the hazard of dropping a label, the risk set contains all dyads of an organization that has previously patented and their labels: 5,967 dyads over 10,908 years, with 4,397 events of dropping a label. Again we also included the nonpatenters in additional analyses; the entire data set contains 21,589 organization-label dyads over 39,359 organization-label-years, with 13,880 events of dropping a label. We analyze spells during the years 1990–2001. Dyads enter the risk set in the first year the organization affiliates with the label. Duration is the time elapsed since the organization first affiliates with the label.

### Hypotheses

We test four hypotheses in this empirical context based on our theoretical propositions. The first and second hypotheses concern the relationship between feature-space positions and label affiliations (prop. 1).

**Hypothesis 1.** A producer's hazard of adding a label increases with its proximity in knowledge space to the cluster of producers associated with the label.

**Hypothesis 2.** A producer's hazard of dropping a label decreases with its proximity in knowledge space to the cluster of producers associated with the label.

The third and fourth hypotheses concern the effects of leniency (prop. 2).

**Hypothesis 3.** Proximity in knowledge space to the producers associated with a constraining label has a stronger positive effect on the hazard of adding that label than does proximity to a more lenient label.

**Hypothesis 4.** Proximity in knowledge space to the producers associated with a constraining label has a stronger negative effect on the hazard of dropping that label than does proximity to a more lenient label.

## Results

### Adding Labels

We analyze the effects of knowledge space proximity on the hazards of adding a label. Table 3 reports estimations on the hazard of adding a label for the independent variables in the analysis (all estimations include controls; full models available upon request). The first model in Table 3 reports tests of H1. This analysis considers dyads of organizations that have patented and the labels that they do *not* already claim. The results support H1: an organization's knowledge-space proximity to a label has a positive effect on the hazard of it adopting this label ( $p < 0.01$  in a z-score test). An organization that is two standard deviations above the mean on proximity to a label is 10 percent more likely to adopt the label, as compared to one at mean proximity, according to this estimate.

Next we analyze whether the effects of proximity depend on leniency. The second model of Table 3 provides tests of H3. The results support this hypothesis: proximity to a constraining label has a strong positive effect on the hazard of adding it (z-score test significant at  $p < 0.01$ ). However, the effect of proximity is smaller and insignificant when the proximate label is lenient. This specification fits better than the one in model 1 (the likelihood-ratio test is significant at  $p < 0.01$ ). This indicates that feature values matter less for affiliation with lenient labels.

We also explored some possible confounding effects. One question might be whether proximity drives the effect, or whether it just reflects the consequences of general exploration in the space. To test against this alternative, we include an effect of the organization's proximities to *all* other labels it does not affiliate with, and to all other labels it does affiliate with (model 3). The results show that general exploration in knowledge space has a positive effect on an organization's propensity to add any label. But the effect of exploration is smaller by an order of magnitude

as compared to the effect of proximity. An organization's proximity to labels it already claims does not have a statistically significant effect on adding a new label. Importantly, our main results persist. Estimations run on all organizations in the press release data, where nonpatenters are assigned knowledge proximity zero, gives similar results (model 4). Results are also similar when we control for leniency using a binary variable that indicates whether the organization affiliates with a high-lenient label, defined using the same cutoff.

### Dropping Labels

The next set of analyses explores the effects of knowledge-space proximity on the hazard of dropping a label. Table 4 reports estimations on the hazard of dropping a label for the independent variables in the analysis (all estimations include controls; full models available upon request). The first model in Table 4 provides a test of H2. This analysis considers dyads of organizations and labels that they affiliate with. The results support H2: the higher an organization's proximity to a label in knowledge space, the less likely it will drop its affiliation with that label ( $p < 0.05$  in a z-score test). This specification also includes an effect of the organization's knowledge space proximity to all other unclaimed labels. This effect is positive. In some estimations (Table 4, models 4 and 5), the effect is positive and significant. Together, these results illustrate the dynamic link between feature space location and label assignments: the closer is an organization to a label it already claims, the less likely it is to drop this label. The closer it is to other labels it does not claim, the more likely it is to drop a claimed label (and adopt the new label to which it has become proximate).

The second model in Table 4 provides tests of H4, concerning the effects of leniency on hazards of dropping a label. The coefficients have the predicted signs but they do not differ significantly according to a chi-square test. Consistent with the hypothesis, proximity to a constraining label has a negative and significant effect on the hazard of dropping (z-score test significant at  $p < 0.01$ ), whereas the effect for lenient labels is weaker and not significant. To explore this effect further, we increased the leniency threshold from 2.7 to 3.0,

**Table 3:** Effects of Proximity and Leniency on the Hazards of Adding a Label (ML Estimates of Piecewise-Continuous Hazard Models)

	Hyp. 1 (Patenters) (1)	Hyp. 3 (Patenters) (2)	Hyp. 3 (Patenters) (3)	Hyp. 3 (All Orgs) (4)
K-space proximity				
Focal label	1.049 <sup>†</sup> (0.243)			
Constraining focal label		1.548 <sup>†</sup> (0.208)	1.470 <sup>†</sup> (0.195)	1.464 <sup>†</sup> (0.199)
Lenient focal label		0.442 (0.368)	0.205 (0.370)	−0.069 (0.324)
Other not claimed label(s)			0.153 <sup>†</sup> (0.025)	0.135 <sup>†</sup> (0.023)
Label(s) org claims			0.063 (0.057)	0.051 (0.053)
Log pseudo-likelihood	−23,769.69	−23,754.75	−23,717.94	−75,284.90
Degrees of freedom	28	29	31	31

*Note:* All specifications include: label controls for leniency, fuzzy density, recent adds, recent drops, and label tenure (since 1990); organizational controls for the number of labels the organization claims, the number of other labels in the industry, a count of the organization's patents, tenure (since 1990), whether it ranked in the Software 500, recently received venture capital financing, and the time since adding any label; year fixed effects and duration pieces.

Patenters: 337,800 org-label dyads over 1,434,096 org-label-years; 5,266 events of adding a label.

All orgs: 1,893,569 org-label dyads over 6,485,521 org-label-years; 15,103 events of adding a label.

\*  $p < 0.05$ ; †  $p < 0.01$

and the results strongly support the hypothesis (model 3). This effect persists when we include a control for an organization's proximity to the other labels it already claims (model 4). An estimation on all organizations in the press-release data (model 5) also shows similar effects. In addition, consistent results hold when we control for leniency using a binary variable.

This analysis highlights the importance of the existential quantification in the definition of leniency: the predictions about leniency hold only provided that one category is *sufficiently* more lenient than another. In this analysis, these effects are evident when the leniency threshold is increased, suggesting that there exists a leniency threshold above which the link between the spaces breaks down. Overall, this analysis provides some support for H4. Together with the results reported here, it provides additional evidence that leniency weakens the relationship

between feature values and label affiliations. It also suggests that feature-space proximity has a stronger influence on the propensity to drop a lenient label as compared to the propensity to adopt a lenient label.

Proximity to *other* labels the organization affiliates with has a negative effect on the hazard of dropping the focal label, and it is significant in analyses of all organizations ( $z$ -score test significant at  $p < 0.01$ ). Again, the size of the effect is substantially smaller than the effect of knowledge-space proximity to the focal label (from the dyad). This indicates that there might be some symbiotic effects in terms of the feature profiles of labels a producer simultaneously claims.

One potential concern is that organizations might generally be more proximate to lenient labels and that this drives the results in both the estimates for both adding and dropping labels. Our data show that the distributions of

**Table 4:** Effects of Proximity and Leniency on the Hazards of Dropping a Label (ML Estimates of Piecewise-Continuous Hazard Models)

	Hyp. 2 (Patenters) (1)	Hyp. 4 (Patenters) (2)	Hyp. 4 (Patenters) (3)	Hyp. 4 (Patenters) (4)	Hyp. 4 (All Orgs) (5)
K-space proximity to:					
Focal label	−0.422* (0.186)				
Constraining focal label		−0.607† (0.212)	−0.716† (0.223)	−0.698† (0.218)	−0.680† (0.220)
Lenient focal label		−0.251 (0.260)	−0.035 (0.180)	−0.045 (0.177)	−0.215 (0.205)
Not claimed label(s)	0.036 (0.020)	0.038 (0.020)	0.038 (0.020)	0.062* (0.031)	0.108† (0.027)
Other label(s) org. claims				−0.035 (0.031)	−0.079* (0.031)
Log pseudo-likelihood	−6,705.53	−6,705.05	−6,702.67	−6,701.61	−24,678.40
Degrees of freedom	29	30	30	31	31

*Note:* All specifications include: label controls for leniency, fuzzy density, recent adds, recent drops, and label tenure (since 1990); organizational controls for the number of labels the organization claims, the number of other labels in the industry, a count of the organization's patents, tenure (since 1990), whether it ranked in the Software 500, recently received venture capital financing, and the time since dropping any label; year fixed effects and duration pieces. Model 3 imposes a higher leniency threshold (see text).

Patenters: 5,967 org-label dyads over 10,908 org-label-years; 4,397 events of dropping a label.

All orgs: 21,589 org-label dyads over 39,359 org-label-years; 13,880 events of dropping a label.

\*  $p < 0.05$ ; †;  $p < 0.01$

knowledge-space proximities to constraining and lenient labels are similar.<sup>26</sup>

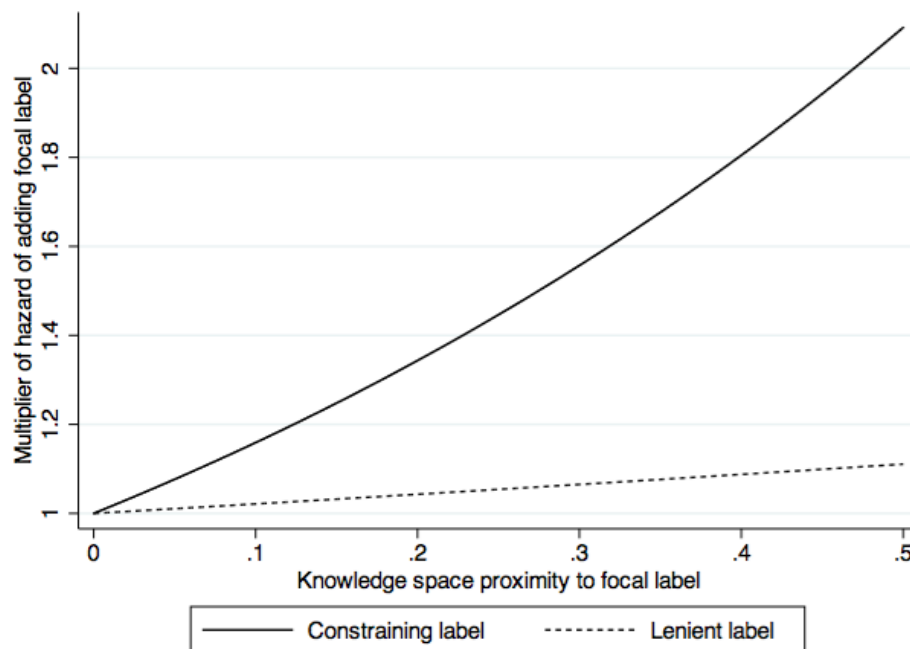
Together these results for dropping labels and those reported for adding labels show how feature-based proximities shape categorical affiliations in a dynamic context. These effects are illustrated in Figures 5 and 6. For constraining labels, there is a strong relationship between an organization's proximity to the label and an increased hazard of adding the label, and decreased hazard of dropping the label. For lenient labels, the link is weaker.

## Additional Tests

We conduct additional tests of these results. Our model proposes a coupled ecology between feature space and label space, where both positions of organizations and meanings of categories shift

over time. Results show that the relative distance in feature space between an organization and category representation affects whether an organization adds or drops the category label. We further explore this by including in the estimations a measure of organization stationarity, which measures the percentage of the organization's current year citations that the organization has previously cited (in the past four years). Organizations that score low on this measure have moved more in knowledge space, whereas those that score high are more stationary. We include stationarity and an interaction between stationarity and knowledge proximity to the focal label as covariates. Results show that neither stationarity nor the interaction have a significant effect on an organization's propensity to add or drop labels. Effects reported here persist (results available upon request). This suggests that the reported effects do not simply reflect an organization's

<sup>26</sup>See Tables 1 and 2.



**Figure 5:** Effects of knowledge proximity on the hazard of adding a label, by leniency (from Table 3, model 3).

movement in feature space toward a stationary category.

We also conduct additional tests to further control for social factors that might affect an organization's propensity to add or drop labels. We control for the number of organizations in a label that have recently received venture-capital financing. Such investments signal that an area is promising and might give managers hope that additional funds will follow. The hypothesized effects are robust to the inclusion of this control.

## Discussion

### Constraint or Accident?

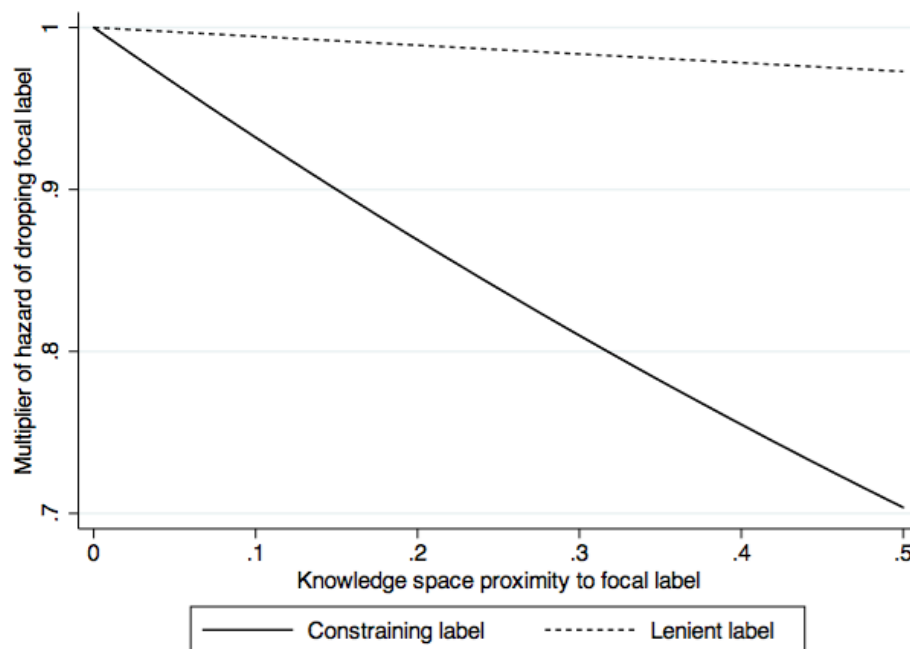
We noted that an observed pattern of label overlap can result from accidental circumstances or from constraint. We suggest that these patterns are not accidental and that they reflect differences in constraint among labels. But if constraint is the source of variation in the strength of categorical boundaries, where does this constraint come from? Here we sketch an extension of the

theory that links category overlap to appeal for producers that change positions in label space.

Suppose that an audience member lacks detailed information about the feature values of two producers but does observe label assignments.<sup>27</sup> Consider two producers that belong to different categories, one constraining and one much more lenient with equal (high) typicality in their assigned category. Suppose that each is assigned to some third label. Recall that our argument implies that constraining categories normally lie further from other categories than do lenient ones (if the difference in leniency exceeds a critical value). Then the definition of the typicality in a label ( $\vartheta$ ) implies that the member of a constraining category generally experiences a greater decline in its typicality its initial category than is the case for the member of a lenient one. In other words, taking on new label assignments generally reduces typicality more for members of constraining categories.

<sup>27</sup>This example, where label assignments are known but feature values are not directly observed, is common in sociological research on classification.





**Figure 6:** Effects of knowledge proximity on the hazard of dropping a label, by leniency (from Table 4, model 4).

This difference matters when the categories involved have positive valuation. This means the intrinsic<sup>28</sup> appeal of a producer to a typical audience member is a positive increasing function of its typicality in the meaning of the relevant label (Hannan et al. 2007).<sup>29</sup> So long as the concepts are positively valued, producers affiliated with constraining labels lose more intrinsic appeal when they span concepts than is the case for those affiliated with more lenient ones.

**Proposition 3.** *Suppose that one producer belongs to one but not the other category and a second producer has the mirror image pattern of memberships and the typicalities are equal. If*

<sup>28</sup>“Intrinsic” in this context refers to degree of fit to the agent’s aesthetics. The theory on which we build holds that it takes engagement of the audience by the producer to convert intrinsic appeal to actual appeal.

<sup>29</sup>Producers can have high (or low) typicality in constraining or lenient categories. The fact that appeal increases with a producer’s typicality in a category need not imply that audiences prefer constraining categories. Previous research shows that venture capitalists, an important audience in the domain of our empirical study, prefer producers in vaguely bounded categories (Pontikes 2012).

*each producer adds the same third label, then the member of the constraining category ends up with lower intrinsic appeal in that label as compared to the member of the more lenient category.*

In this sense, members of constraining categories have more to lose when they adopt feature values that cause them to get assigned some additional category membership(s). This cost is the source of the constraint.

## Dynamics of Leniency

We proposed that proposition 2 implies that lenient labels tend to become more lenient and constraining labels get more constraining. Here we describe how to derive these implications. We use an illustration of a domain with one constraining label and three lenient ones in Figure 7. (For simplicity of representation we use a Euclidean space in this illustration.) In each case, the shaded circle represents a schema (the set of feature-value profiles that lie in these circles are the schemas). The outer circles indicate the locations of the producers assigned the category label.

Category A is constraining because its members more closely fit the schema (the outer circle is closer to the shaded one than is the case for the other categories) and because its membership does not overlap the memberships of the others. Categories B, C, and D are fuzzier (many members far from the schemas) and all three pairs overlap.

Now consider three producers whose feature values place them close to the producers assigned one of the labels. As drawn, producers  $x_1$  and  $x_2$  stand at the same distance from the schema for A. But  $x_1$  lies further from the memberships of the other categories. Thus, according to the theory,  $x_1$  has higher typicality in A than does  $x_2$  and therefore has a higher hazard of being labeled as an A. This means that the expected waiting time to labeled membership is lower for  $x_1$ . In typical cases,  $x_1$  will be labeled an A before  $x_2$ . If this occurs, then the membership of A drifts to the left, away from the memberships of the other categories.

Next compare the situations for  $x_3$  and  $x_4$ . As the illustrative figure is drawn,  $x_3$  lies outside the membership of each category. In this respect it is similar to  $x_1$  and  $x_2$ , and nearly the same reasoning applies. This producer ( $x_3$ ) is slightly further away from the schema for D than is  $x_4$ ; and  $x_3$  is less proximate to the memberships of all categories than is  $x_4$ . So, there is a high likelihood that  $x_3$  will gain the label D. If this happens, it will cause the membership of D to move to the right, away from the other categories, similar to the situation with category A. But  $x_4$  is much closer to the schemas for B and C (and A for that matter) than is  $x_3$ . It has a high hazard of becoming labeled as a member of B. If such an event does occur, this creates a membership overlap where one did not exist (as the figure is drawn). Obviously this new overlap increases the leniency of B and D. The situation is more complicated than in the analysis of A, and the net result depends on the flow of the stochastic events. But there is no clear pattern as was the case for the constraining category.

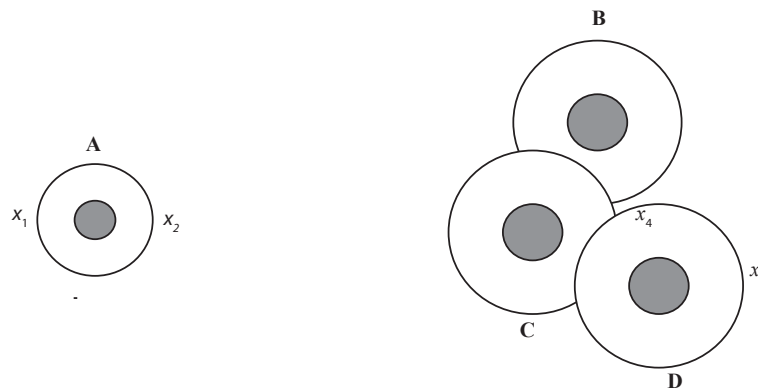
We estimated growth models (at the category level) for leniency, specifications in which leniency in a year depends linearly on leniency in the previous year, the fuzzy density of the category, its tenure, and yearly fixed effects. With either category fixed effects or category random effects,

we find a positive and significant effect of leniency on the growth in leniency. In other words, the higher is a category's leniency, the higher is its growth rate in leniency.

These dynamics likely feed back to the schemas. For instance, as a constraining category drifts away from the lenient ones, what is typical of this category is also shifting away. If, as we expect, audience members try to capture typicality in forming schemas, then the schema will also drift away from the others. This makes the initially constraining category even more distinctive.

If a lenient label remains lenient or becomes more so, then the broad (or broadening) overlap lowers the chance that the audience reaches a consensus about the association of a schema with the label. Put differently, increased leniency diminishes agreement about the meanings. Pockets might develop where audiences have experience with organizations that have one type of overlap with a lenient label, and they might develop a schema with respect to that subset of activities. This likely conflicts with the schema developed by another audience exposed to other organizations that are members of the lenient label but have a different set of overlaps. This might lead to a lack of consensus about the label's schema for relevant audiences. Hence lenient labels might evolve to have ambiguous and uncertain meanings. At best their schemas are defined at a very general level.

This dynamic is reflected in the trajectory of a number of prominent lenient categories in our analysis. For instance, this is the process James Kobielski refers to in his "What's not BI?" blog, when he states that "almost every data management technology has been swept into BI's gravitational orbit at one time or another by somebody somewhere" (Kobielski 2010). Gartner comes to a similar conclusion in their 2006 magic quadrant report on <business intelligence>, noting that <BI> is increasing in scope to encompass more types of users, accessing more information sources and a wider array of applications. <BI> becomes even more lenient over time. In Gartner's (2006) report, they define <business intelligence> as "the mission of BI . . . is the access to and analysis of quantitative information sources to deliver insight that empowers decision makers" (Schlegel et al. 2006). Their 2008 report provides an even less constraining definition that does not specify that the platform should connect the user with



**Figure 7:** Illustration of the dynamic implications of the theory.

disparate information sources: “BI platforms enable users to build applications that help organizations learn and understand their business” (Richardson et al. 2008).

A similar trend is apparent in the evolution of <enterprise resource planning (ERP)> and <customer relationship management (CRM)>. Initially lenient categories, both have increased in leniency over time to the point that they no longer even pose the constraint that category members produce software. Gartner’s (2010) report on <ERP> describes the market category as having evolved from a technology product to a strategy:

In the original definition, ERP systems’ functionality normally covers the following areas: finance and accounting . . . purchasing, human resource management, sales or customer order management, and operations management. However, Gartner now defines ERP in a broader sense as “a technology strategy in which operational business transactions are linked to financial transactions, specifically general ledger transaction.” (Hestermann et al. 2010)

The editors at *CRM* magazine similarly describe <CRM> as having evolved from a type of software to a business philosophy:

CRM, or Customer Relationship Management, is a company-wide business strategy designed to reduce costs and

increase profitability by solidifying customer loyalty. . . Once thought of as a type of software, CRM has evolved into a customer-centric philosophy that must permeate an entire organization. (*CRM Magazine* 21 February 2010).

These examples reflect the dynamic we described: atypical members are more likely to be assigned membership in a lenient label, leading to increased leniency and a broadening of the label’s social meaning.

## Conclusion

In this article we argue that social categorization is governed by an ecological dynamic in two planes: feature space and label space. When an actor changes its feature values or drops or adds a label, these changes not only affect the focal actor but also contribute to changing the meanings of social categories. This, in turn, can lead to cascades of changes in classification, with entities shifting positions in both spaces in response to movements of others. Such uncoordinated movements might seem to lead to chaos. On the contrary, we suggest that this dynamic can perpetuate a link between labels and sets of feature values, if actors pay attention to the positions of others and use labels accordingly. At the same time, it is possible for categories to become lenient in label space, with high degrees of overlap and vague boundaries. This creates an ecology within label space, with proximities to *other* la-

bels disrupting a clear relationship between sets of feature values and perceived typicality in the category. The result is that leniency weakens the relationship between feature-based similarity and categorical assignments.

To investigate these ideas, we built a formal model relating positions in the two spaces. Our theory implies that an object's proximity to labeled clusters in feature space influences the adding or dropping of label affiliations in label space. Furthermore, it predicts that for lenient labels, the link between feature values and labels weakens. The importance of expressing these ideas in formal language comes through in our analysis. It turns out to be tricky to formally derive these conclusions, which suggests that they might not be as straightforward as would appear if expressed in natural language. Furthermore, our analysis highlights the assumptions necessary to arrive at our conclusions.

Our empirical analysis tests the theory's implications in the software industry. We use data on organizations' patents to represent positions in feature space and affiliations with market labels from press releases to represent positions in label space. The results support the new theoretical propositions. We find evidence that when producers are proximate in knowledge space to a label they affiliate with, they are less likely to drop it. If they lie close in knowledge space to labels they do not affiliate with, they are more likely to drop any label and add the more proximate one. We also find that feature-space proximity to lenient labels has a weaker effect on adding a label, as compared to proximity to constraining labels.

In deciding whether to affiliate with a label, agents compare their own feature values with those that already claim the label. This means that the specific features associated with labels can (and do) change over time. Even so, a link between features and labels can persist. The relationship between features and labels is reinforced simply when agents observe how others are classified and respond in kind. It is informative to note what we do not assume: it is not necessary to have strong mechanisms of boundary enforcement or third parties such as critics and intermediaries to retain a link between features and labels. It is enough for agents to infer typicality from feature-based similarities to a labeled cluster and to label based on typicality.

At the same time, our study shows that increasing leniency weakens the relationship between features and label affiliations, resulting in a partial decoupling between the two spaces for lenient labels. With informal classification, both constraint and leniency get reinforced as actors independently navigate feature space and label space. The result is that classification evolves to contain a mix of categories: some constraining, some lenient.

How labels *become* lenient is a different matter. Leniency likely emerges when atypical entities frequently are assigned membership in a label and do not face censure. This suggests that a lack of boundary enforcement can foster the emergence of lenient labels in a classification. But importantly, these elements are not necessary to propagate lenient labels once they exist. This means that once a label becomes lenient, it is likely to remain that way, even with an audience that attends to the links between labels and feature values.

This study also indicates that actors compare positions in feature space and label space *relative* to the positions of other actors in the domain. It is proximity to objects that are affiliated with a label—not necessarily fixed positions in feature space—that drives labeling. In this way, a producer can become more or less proximate to a labeled cluster if the cluster moves in feature space, even if the producer does not move. We believe this captures an interesting facet of social classification: that feature values associated with existing categories can (and do) change, and that changes affect perceptions of the category.

In summary, this article explores dynamics in market classification that arise when producers compare their feature positions to those of competitors and use these comparisons to inform market-label affiliation. Producers evaluate how similar their organizations are to others when deciding whether to adopt or drop a market label. This action does not go unnoticed by others who are similarly comparing feature positions and label affiliations. Thus emerges an ecology of social categories, where positions in feature space and label space change, but a link between them persists, resulting in a dynamic and meaningful system of classification.

## References

- Alcácer, Juan and Michelle Gittleman. 2006. "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations." *The Review of Economics and Statistics* 88:774–779. <http://dx.doi.org/10.1162/rest.88.4.774>.
- Berger, Jonah and Gaël Le Mens. 2009. "How Adoption Speed Affects the Abandonment of Cultural Tastes." *Proceedings of the National Academy of Sciences* 106:8146–8150. <http://dx.doi.org/10.1073/pnas.0812647106>.
- Berger, Peter L. and Thomas Luckmann. 1967. *The Social Construction of Reality*. New York: Anchor Books.
- Black, Max. 1937. "Vagueness: An Exercise in Logical Analysis." *Philosophy of Science* 4:427–455. <http://dx.doi.org/10.1086/286476>.
- Bowker, Geoffrey C. and Susan Leigh Star. 2000. *Sorting Things Out: Classification and its Consequences*. Cambridge: MIT Press.
- Clemens, Elisabeth and James Cook. 1999. "Politics and Institutionalism: Explaining Durability and Change." *Annual Review of Sociology* 25:441–66. <http://dx.doi.org/10.1146/annurev.soc.25.1.441>.
- Cockburn, Iain M. and Megan J. MacGarvie. 2009. "Patents, Thickets and the Financing of Early-Stage Firms: Evidence from the Software Industry." *Journal of Economics and Management Strategy* 18:729–773. <http://dx.doi.org/10.1111/j.1530-9134.2009.00228.x>.
- Cohen, Julie and Mark Lemley. 2001. "Patent Scope and Innovation in the Software Industry." *California Law Review* 89:1–57. <http://dx.doi.org/10.2307/3481172>.
- DiMaggio, Paul J. 1997. "Culture and Cognition." *Annual Review of Sociology* 23:263–287. <http://dx.doi.org/10.1146/annurev.soc.23.1.263>.
- Durkheim, Émile. 1931. *The Elementary Forms of Religious Life*. Oxford: Oxford University Press.
- CRM Magazine. 21 February 2010. "What is CRM?" <http://www.destinationcrm.com/Articles/CRM-News/Daily-News/What-Is-CRM-46033.aspx>.
- Gärdenfors, Peter. 2004. *Conceptual Spaces: The Geometry of Thought*. Cambridge: MIT Press.
- Granqvist, Nina, Stine Grodal, and Jennifer Wooley. 2013. "Hedging Your Bets: Explaining Executives' Market Labeling Strategies in Nanotechnology." *Organization Science* 24:395–413. <http://dx.doi.org/10.1287/orsc.1120.0748>.
- Grodal, Stine. 2011. "Boundary Labeling During Field Emergence: Inclusion and Exclusion in Nanotechnology." Working paper, Boston University School of Management.
- Hahn, Ulrike, Nick Chater, and Lucy B. Richardson. 2003. "Similarity as Transformation." *Cognition* 87:1–32. [http://dx.doi.org/10.1016/S0010-0277\(02\)00184-1](http://dx.doi.org/10.1016/S0010-0277(02)00184-1).
- Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg. 2001. *The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools*. Cambridge, MA: National Bureau of Economic Research. <http://dx.doi.org/10.3386/w8498>.
- Hampton, James A. 2007. "Typicality, Graded Membership, and Vagueness." *Cognitive Science* 31:355–383. <http://dx.doi.org/10.1080/15326900701326402>.
- Hannan, Michael T., László Pólos, and Glenn R. Carroll. 2003. "Cascading Organizational Change." *Organization Science* 14:463–482. <http://dx.doi.org/10.1287/orsc.14.5.463.16763>.
- Hannan, Michael T., László Pólos, and Glenn R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes and Ecologies*. Princeton: Princeton University Press.
- Hestermann, Christian, Chris Pang, and Nigel Montgomery. 2010. "Magic Quadrant for ERP for Product-Centric Midmarket Companies." Gartner.



- Hsu, Greta. 2006. "Jacks of All Trades and Masters of None: Audiences' Reactions to Spanning Genres in Feature Film Production." *Administrative Science Quarterly* 51:420–450. <http://dx.doi.org/10.2189/asqu.51.3.420>.
- Hsu, Greta and Kimberly D. Elsbach. 2013. "Explaining Variation in Organizational Identity Categorization." *Organization Science* 24:996–1013. <http://dx.doi.org/10.1287/orsc.1120.0779>.
- Hsu, Greta, Michael T. Hannan, and Özgecan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests." *American Sociological Review* 74:150–69. <http://dx.doi.org/10.1177/000312240907400108>.
- Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." *American Economic Review* 76:984–1001.
- Jaffe, Adam B. 1989. "Real Effects of Academic Research." *American Economic Review* 79:957–970.
- Keefe, Rosanna and Peter Smith (eds.). 1997. *Vagueness: A Reader*. Cambridge Mass.: MIT Press.
- Kobielus, James. 2010. "What's Not BI? Oh, Don't Get Me Started...Oops Too Late...Here Goes..." Forrester Blogs, [http://blogs.forrester.com/james\\_kobielus/10-04-30](http://blogs.forrester.com/james_kobielus/10-04-30).
- Koçak, Özgecan, Michael T. Hannan, and Greta Hsu. 2014. "Emergence of Market Orders: Audience Interaction and Vanguard Influence." *Organization Studies* 35:765–790. <http://dx.doi.org/10.1177/0170840613511751>.
- Kovács, Balázs and Michael T. Hannan. 2011. "Category Spanning, Distance, and Appeal." Research Report 2081, Stanford Graduate School of Business.
- Kovács, Balázs and Rebeka Johnson. 2014. "Contrasting Alternative Explanations for the Negative Effect of Category Spanning: A Study of Restaurant Reviews and Menus in San Francisco." *Strategic Organization* 12:7–37. <http://dx.doi.org/10.1177/1476127013502465>.
- Leung, Ming and Amanda Sharkey. 2014. "Out of Sight, Out of Mind: Evidence of Perceptual Factors in the Multiple-Category Discount." *Organization Science* 25:171–184. <http://dx.doi.org/10.1287/orsc.2013.0828>.
- Lieberson, Stanley. 2000. *A Matter of Taste: How Names, Fashions, and Culture Change*. New Haven CT: Yale University Press.
- Mann, Ronald J. 2005. "Do Patents Facilitate Financing in the Software Industry?" *Texas Law Review* 83:1–71.
- Mann, Ronald J. and Thomas W. Sager. 2007. "Patents, Venture Capital, and Software Start-ups." *Research Policy* 36:193–208. <http://dx.doi.org/10.1016/j.respol.2006.10.002>.
- McKendrick, David, Jonathan Jaffee, Glenn Carroll, and Olga Khessnia. 2003. "In the Bud? Disk Array Producers as a (Possibly) Emergent Organizational Form." *Administrative Science Quarterly* 48:60–93. <http://dx.doi.org/10.2307/3556619>.
- Meyer, John and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83:340–63. <http://dx.doi.org/10.1086/226550>.
- Meyersohn, Rolf and Elihu Katz. 1957. "Notes on a Natural History of Fads." *American Journal of Sociology* 62:594–601. <http://dx.doi.org/10.1086/222108>.
- Murphy, Gregory L. 2002. *The Big Book of Concepts*. Cambridge: MIT Press.
- Negro, Giacomo, Michael T. Hannan, and Haya-greeva Rao. 2010. "Categorical Contrast and Audience Appeal: Niche Width and Critical Success in Winemaking." *Industrial and Corporate Change* 19:1397–1425. <http://dx.doi.org/10.1093/icc/dtq003>.
- Negro, Giacomo and Ming D. Leung. 2013. "'Actual' and Perceptual Effects of Category Spanning." *Organization Science* 24:684–696. <http://dx.doi.org/10.1287/orsc.1120.0764>.
- Podolny, Joel M. and Toby E. Stuart. 1995. "A Role-Based Ecology of Technological Change." *American Journal of Sociology* 100:1224–1260. <http://dx.doi.org/10.1086/230637>.



- Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan. 1996. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984–1991." *American Journal of Sociology* 102:659–689. <http://dx.doi.org/10.1086/230994>.
- Pólos, László and Michael T. Hannan. 2004. "A Logic for Theories in Flux: A Model-Theoretic Approach." *Logique et Analyse* 47:85–121.
- Pontikes, Elizabeth G. 2008. *Fitting In or Starting New? An Analysis of Invention, Constraint, Contrast, and the Emergence of New Categories in the Software Industry*. Ph.D. thesis, Stanford University, Stanford, CA.
- Pontikes, Elizabeth G. 2012. "Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audience Evaluations." *Administrative Science Quarterly* 57:81–118. <http://dx.doi.org/10.1177/0001839212446689>.
- Porac, Joseph F., Howard Thomas, Fiona Wilson, Douglas Paton, and Alaina Kanfer. 1995. "Rivalry and the Industry Model of Scottish Knitwear Producers." *Administrative Science Quarterly* 20:203–227. <http://dx.doi.org/10.2307/2393636>.
- Rao, Hayagreeva, Phillipe Monin, and Rodolphe Durand. 2003. "Institutional Change in Toque Ville: Nouvelle Cuisine as an Identity Movement in French Gastronomy." *American Journal of Sociology* 108:795–843. <http://dx.doi.org/10.1086/367917>.
- Richardson, James, Kurt Schlegel, Bill Hostmann, and Neil McMurchy. 2008. "Magic Quadrant for Business Intelligence Platforms, 2008." Gartner.
- Romanelli, Elaine. 1991. "The Evolution of New Organizational Forms." *Annual Review of Sociology* 17:79–103. <http://dx.doi.org/10.1146/annurev.soc.17.080191.000455>.
- Rosch, Eleanor H. and Barbara Lloyd (eds.). 1978. *Principles of Categorization*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Rosch, Eleanor H. and Carolyn B. Mervis. 1975. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7:573–605. [http://dx.doi.org/10.1016/0010-0285\(75\)90024--9](http://dx.doi.org/10.1016/0010-0285(75)90024--9).
- Ruef, Martin. 2000. "The Emergence of Organizational Forms: A Community Ecology Approach." *American Journal of Sociology* 106:658–714. <http://dx.doi.org/10.1086/318963>.
- Ruef, Martin and Kelly Patterson. 2009. "Credit and Classification: Defining Industry Boundaries in 19th Century America." *Administrative Science Quarterly* 54:486–520. <http://dx.doi.org/10.2189/asqu.2009.54.3.486>.
- Schlegel, Kurt, Bill Hostmann, Andreas Bitterer, and Betsy Burton. 2006. "Magic Quadrant for Business Intelligence Platforms, 1Q06." Gartner.
- Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science* 237:1317–1323. <http://dx.doi.org/10.1126/science.3629243>.
- Smith, Edward E. and Douglas L. Medin. 1981. *Categories and Concepts*. Cambridge, MA: Harvard University Press. <http://dx.doi.org/10.4159/harvard.97806748666270>.
- Soltes, Eugene. 2009. "News Dissemination and the Impact of the Business Press." Dissertation, The University of Chicago.
- Sørensen, Jesper B. and Toby E. Stuart. 2000. "Aging, Obsolescence, and Organizational Innovation." *Administrative Science Quarterly* 45:81–112. <http://dx.doi.org/10.2307/2666980>.
- Strang, David and Sarah A. Soule. 1998. "Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills." *Annual Review of Sociology* 24:265–90. <http://dx.doi.org/10.1146/annurev.soc.24.1.265>.
- Stuart, Toby E. and Joel M. Podolny. 1996. "Local Search and the Evolution of Technological Capabilities." *Strategic Management Journal* 17:21–38. <http://dx.doi.org/10.1002/smj.4250171004>.
- Tversky, Amos. 1977. "Features of Similarity." *Psychological Review* 84:327–351. <http://dx.doi.org/10.1037/0033-295X.84.4.327>.

- Tversky, Amos and Itamar Gati. 1978. "Studies of Similarity." In *Cognition and Categorization*, edited by Eleanor Rosch and Barbara Lloyd. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- van Demeter, Kees. 2010. *Not Exactly: In Praise of Vagueness*. Oxford: Oxford University Press.
- Verheyen, Steven, James A. Hampton, and Gert Storms. 2010. "A Probabilistic Threshold Model: Analyzing Semantic Categorization Data With the Rasch Model." *Acta Psychologica* 135:216–225. <http://dx.doi.org/10.1016/j.actpsy.2010.07.002>.
- White, Harrison. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 87:517–547. <http://dx.doi.org/10.1086/227495>.
- Widdows, Dominic. 2004. *Geometry and Meaning*. Stanford CA: CSLI Publications.
- Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Illegitimacy Discount." *American Journal of Sociology* 104:1398–1438. <http://dx.doi.org/10.1086/210178>.

**Acknowledgements:** We appreciate the extremely helpful suggestions of Greta Hsu, Balázs Kovács, Gaël Le Mens, László Pólos, Amanda Sharkey, and Chris Yenkey. This study was supported by the Chicago-Booth Graduate School of Business, the Stanford Graduate School of Business, and the Charles E. Merrill Faculty Research Fund at the University of Chicago Booth School of Business, and by the John Osterweis and Barbara Ravizzi Faculty Fellowship at the Stanford Graduate School of Business.

**Elizabeth G. Pontikes:** University of Chicago.  
E-mail: [elizabeth.pontikes@chicagobooth.edu](mailto:elizabeth.pontikes@chicagobooth.edu).

**Michael T. Hannan:** Stanford University.  
E-mail: [hannan@stanford.edu](mailto:hannan@stanford.edu).